

PCT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents
 United States Patent and Trademark
 Office
 Box PCT
 Washington, D.C.20231
 ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 28 June 2000 (28.06.00)	
International application No. PCT/EP99/09833	Applicant's or agent's file reference SCB/53731/00
International filing date (day/month/year) 06 December 1999 (06.12.99)	Priority date (day/month/year) 04 December 1998 (04.12.98)
Applicant CONTRERAS, Roland, Henri et al	

1. The designated Office is hereby notified of its election made:



in the demand filed with the International Preliminary Examining Authority on:

12 May 2000 (12.05.00)



in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was

was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer A. Karkachi Telephone No.: (41-22) 338.83.38
---	---

PCT COOPERATION TREATY

PCT

From the INTERNATIONAL BUREAU

NOTIFICATION OF THE RECORDING
OF A CHANGE(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

To:

BOULT WADE TENNANT
27 Furnival Street
London EC4A 1PQ
ROYAUME-UNI

Date of mailing (day/month/year) 08 June 2000 (08.06.00)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference SCB/53731/00	
International application No. PCT/EP99/09833	International filing date (day/month/year) 06 December 1999 (06.12.99)

1. The following indications appeared on record concerning:

☒ the applicant ☒ the inventor ☐ the agent ☐ the common representative

Name and Address

VIALARD, Jorge, Edwards
Janssen Pharmaceutica N.V.
Turnhoutseweg 30
B-2340 Beerse
Belgium

State of Nationality

CA

State of Residence

BE

Telephone No.

Facsimile No.

Teleprinter No.

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person ☒ the name ☐ the address ☐ the nationality ☐ the residence

Name and Address

VIALARD, Jorge, Eduardo
Janssen Pharmaceutica N.V.
Turnhoutseweg 30
B-2340 Beerse
Belgium

State of Nationality

CA

State of Residence

BE

Telephone No.

Facsimile No.

Teleprinter No.

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

<input checked="" type="checkbox"/> the receiving Office	<input checked="" type="checkbox"/> the designated Offices concerned
<input checked="" type="checkbox"/> the International Searching Authority	<input type="checkbox"/> the elected Offices concerned
<input type="checkbox"/> the International Preliminary Examining Authority	<input type="checkbox"/> other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

Athina Nickitas-Etienne

Telephone No.: (41-22) 338.83.38

F. ENT COOPERATION TREA

PCT

NOTIFICATION OF THE RECORDING
OF A CHANGE(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

From the INTERNATIONAL BUREAU

To:

Baldock, Sharon, Claire
Boult Wade Tennant
Verulam Gardens
70 Gray's Inn Road
London WC1X 8BT
ROYAUME-UNI

Date of mailing (day/month/year) 28 June 2000 (28.06.00)	
Applicant's or agent's file reference SCB/53731/00	IMPORTANT NOTIFICATION
International application No. PCT/EP99/09833	International filing date (day/month/year) 06 December 1999 (06.12.99)

1. The following indications appeared on record concerning:

☐ the applicant ☐ the inventor ☒ the agent ☐ the common representative

Name and Address BOULT WADE TENNANT 27 Furnival Street London EC4A 1PQ United Kingdom	State of Nationality	State of Residence
	Telephone No. +44 020 7430 7500	
	Facsimile No. +44 020 7831 1768	
	Teleprinter No.	

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☒ the person ☐ the name ☒ the address ☐ the nationality ☐ the residence

Name and Address Baldock, Sharon, Claire Boult Wade Tennant Verulam Gardens 70 Gray's Inn Road London WC1X 8BT United Kingdom	State of Nationality	State of Residence
	Telephone No. +44 020 7430 7500	
	Facsimile No. +44 020 7831 1768	
	Teleprinter No.	

3. Further observations, if necessary:

The new agent's address on the Demand has been considered as a change under Rule 92bis. In case of disagreement, the International Bureau should be notified immediately.

4. A copy of this notification has been sent to:

<input checked="" type="checkbox"/> the receiving Office	<input type="checkbox"/> the designated Offices concerned
<input type="checkbox"/> the International Searching Authority	<input checked="" type="checkbox"/> the elected Offices concerned
<input checked="" type="checkbox"/> the International Preliminary Examining Authority	<input type="checkbox"/> other:

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer A. Karkachi Telephone No.: (41-22) 338.83.38
---	---

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference SCB/53731/00	FOR FURTHER ACTION see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No: PCT/EP 99/ 09833	International filing date (day/month/year) 06/12/1999	(Earliest) Priority Date (day/month/year) 04/12/1998
Applicant JANSSEN PHARMACEUTICA N.V. et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

- a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

☐ the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

- b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

☒ contained in the international application in written form.

☒ filed together with the international application in computer readable form.

☐ furnished subsequently to this Authority in written form.

☐ furnished subsequently to this Authority in computer readable form.

☐ the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

☐ the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. ☒ **Certain claims were found unsearchable** (See Box I).

3. ☒ **Unity of invention is lacking** (see Box II).

4. With regard to the title,

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the abstract,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the drawings to be published with the abstract is Figure No.

☐ as suggested by the applicant.

☐ because the applicant failed to suggest a figure.

☐ because this figure better characterizes the invention.

☒ **None of the figures.**

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 99/09833

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 25-28
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-28 and 32-34, all partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 25-28

Present claims 25-28 relate to a compound defined by reference to a desirable characteristic or property, namely its identifiability by the method of claim 24.

The application does not provide support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search for these claims is impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: Invention 1: claims 1-28 and 32-34, all partially

Nucleic acid molecule comprising seq.ID.1 or capable of hybridizing thereto and antisense thereof, cell containing said nucleic acid, polypeptide of seq.ID.10 encoded by said nucleic acid, expression vector comprising said nucleic acid, antibody against said peptide, use of said nucleic acid or protein for preparation of medicament or pharmaceutical composition, *C. albicans* cell comprising an induced mutation in said DNA sequence, oligonucleotides comprising 10-120 nt of said nucleic acid sequence, and method for identifying compounds which modulate expression of said nucleic acid.

2. Claims: Inventions 2-9: claims 1-28 and 32-34, all partially and as applicable

As invention 1, but limited to the respective nucleic acid sequences 2-9, and polypeptide sequences corresponding thereto in as far as they are provided (see table 1 of the description).

3. Claims: Invention 10: claim 29-31

Method for identifying DNA sequences from a cell or organism, which encode polypeptides which are critical for growth and survival for said cell or organism, comprising screening a library of nucleic acids using a vector that either integrates into the genome of said cell or organism, or that permits expression of antisense RNA, and selecting growth-impaired cells or organisms.

INTERNATIONAL SEARCH REPORT

International Application No.
PC 99/09833

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/31 C07K14/40 G01N33/50 C12Q1/68 A61K31/7088
A61K38/16 C07K16/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>✓ DATABASE EMBL - EMGSS12 [Online] Entry/Acc.no. B89355, 19 March 1998 (1998-03-19) ADAMS, M. D. ET AL.: "RPCI11-27L12.TPB RPCI-11 Homo sapiens genomic clone RPCI-11-27L12, genomic survey sequence." XP002139078 the whole document</p> <p>---</p>	33,34
A	<p>✓ WO 97 36925 A (SCRIPTGEN PHARM INC ;HARVARD COLLEGE (US)) 9 October 1997 (1997-10-09) the whole document</p> <p>---</p>	
A	<p>✓ WO 97 37230 A (SCRIPTGEN PHARM INC ;HARVARD COLLEGE (US)) 9 October 1997 (1997-10-09) the whole document</p> <p>---</p>	
	--- -/-	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

30 May 2000

Date of mailing of the international search report

23.08.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Smalt, R

INTERNATIONAL SEARCH REPORT

International Application No
PC 99/09833

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>WO 96 36707 A (UNIV ROMA ;IST SUPERIORE SANITA (IT); CASSONE ANTONIO (IT); VALLE) 21 November 1996 (1996-11-21) the whole document</p> <p>---</p>	
A	<p>DALY S ET AL: "Isolation and characterization of a gene encoding alpha-tubulin from Candida albicans" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES,GB,ELSEVIER SCIENCE PUBLISHERS, BARKING, vol. 187, no. 2, 18 March 1997 (1997-03-18), pages 151-158, XP004093273 ISSN: 0378-1119 the whole document</p> <p>-----</p>	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

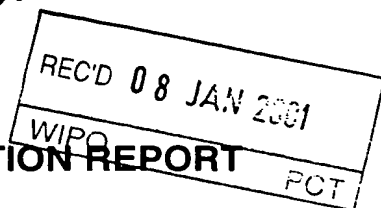
PCT/JP 99/09833

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9736925 A	09-10-1997	CA 2250129 A EP 0904289 A	09-10-1997 31-03-1999
WO 9737230 A	09-10-1997	US 5863762 A CA 2250121 A EP 0894269 A	26-01-1999 09-10-1997 03-02-1999
WO 9636707 A	21-11-1996	IT RM950314 A AU 5777696 A EP 0826040 A	18-11-1996 29-11-1996 04-03-1998

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)




15

Applicant's or agent's file reference SCB/53731/00	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/EP99/09833	International filing date (day/month/year) 06/12/1999	Priority date (day/month/year) 04/12/1998
International Patent Classification (IPC) or national classification and IPC C12N15/31		
Applicant JANSSEN PHARMACEUTICA N.V. et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 8 sheets, including this cover sheet.
- ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).
- These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☒ Certain defects in the international application
- VIII ☒ Certain observations on the international application

Date of submission of the demand 12/05/2000	Date of completion of this report 04.01.2001
Name and mailing address of the international preliminary examining authority:  European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Chavanne, F Telephone No. +49 89 2399 8399



INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP99/09833

I. Basis of the report

1. This report has been drawn on the basis of *(substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments (Rules 70.16 and 70.17).)*:

Description, pages:

1-36 as originally filed

Claims, No.:

1-34 as originally filed

Drawings, No.:

1-31 as originally filed

Sequence listing part of the description, pages:

1-16, as originally filed

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language: , which is:

- ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☒ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP99/09833

- ☐ the description, pages:
- ☐ the claims, Nos.:
- ☐ the drawings, sheets:

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

(Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)

6. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

- ☐ the entire international application.
- ☒ claims Nos. 25-28, completely; 1-24 and 32-34, partially.

because:

- ☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):
- ☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. are so unclear that no meaningful opinion could be formed (*specify*):
- ☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.
- ☒ no international search report has been established for the said claims Nos. 25-28, completely; 1-24 and 32-34, partially.

2. A meaningful international preliminary examination report cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

- ☐ the written form has not been furnished or does not comply with the standard.
- ☐ the computer readable form has not been furnished or does not comply with the standard.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability;

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP99/09833

citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	4, 6, 8, 10-32
	No:	Claims	1-3, 5, 7, 9, 33, 34
Inventive step (IS)	Yes:	Claims	4, 6, 8, 10-32
	No:	Claims	1-3, 5, 7, 9, 33, 34
Industrial applicability (IA)	Yes:	Claims	1-34
	No:	Claims	

2. Citations and explanations see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:
see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:
see separate sheet

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. An objection of lack of unity has been raised by the Searching Authority. Since the Applicant did not pay the requested additional search fees, the International Search Report has been restricted to the invention mentioned as invention 1 covered by claims 1-28 and 32-34, all partially. Moreover, claims 25-28 were found to be unsearchable. Consequently, the present examination has been performed on the searched claims (claims 1-24 and 32-34 partially).

V. Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Reference is made to the following documents:

D1: Database-EMBL

Accession Nr.: B89355

D2: Gene

Vol. 187, No. 2, pp. 151.158, 1997

2. The subject-matter of claims 1-5 refers to a nucleic acid molecule comprising the sequence of SEQ ID No.1. Because of the expression "comprising", the claimed nucleic acid molecule is not necessarily limited to the sequence of SEQ ID No.1, but may also contain other sequences. The genomic DNA of *Candida albicans*, which has already been isolated, is a DNA molecule which comprises the sequence of SEQ ID No.1, and thus, falls within the scope of claims 1-3 and 5. Claims 1-3 try to further characterise the claimed nucleic acid molecule in that it encodes a polypeptide which is critical for survival and growth of the yeast *Candida albicans*. However, in this connection, it is pointed out that as a general rule, the elucidation of a novel feature or property of a known product is not able to reinstate its novelty. Thus, claims 1-3 and 5 lack novelty. Claim 3 further lacks novelty in that a fragment of a sequence includes also e.g. single nucleotides. Since genes of *C. albicans* have been isolated and characterised (see e.g. D2,

which describes the isolation and characterisation of a gene encoding α -tubulin from *C. albicans*), nucleic acid molecules hybridising to the genomic DNA of *C. albicans* are known in the art. Thus, claim 7 lacks novelty.

Eukaryotic cells containing the genomic DNA of *C. albicans*, i.e. the cells of *C. albicans*, are known in the art. Thus, the subject-matter of claim 9 is not novel. Therefore, claims 1-3, 5, 7 and 9 do not meet the requirements of Article 33(2) PCT.

3. D1 describes an oligonucleotide comprising a fragment of 13 contiguous nucleotides of the sequence of SEQ ID No.1.
Thus, in view of D1, the subject-matter of claims 33 and 34 is not novel.
Therefore, claims 33 and 34 do not meet the requirements of Article 33(2) PCT.
4. The subject-matter of claims 4, 6, 8 and 10-32 is neither specifically disclosed nor suggested in the prior art. Thus, claims 4, 6, 8 and 10-32 appear to be novel and inventive.

VII. Certain defects in the international application

1. Independent claims 1-3 disclose a nucleic acid molecule comprising the sequence of SEQ ID No.1. As a consequence of the partial search (see item I of the present communication) present claims 1 and 2 refer to the exact same subject-matter, and the scope of claim 3 is broader than the scope of claims 1 and 2 in that the scope of claim 3 further encompasses fragments or derivatives of said nucleic acid molecule. Thus, it appears appropriate to amend said claims by defining the relevant subject-matter in terms of one single independent claim followed by dependent claims covering the optional features (Rule 6.4 PCT).
2. Claim 10 refers to the polypeptide sequences of SEQ ID No.14. It is probably the result of a typing error since the polypeptide encoded by the nucleic acid molecule of SEQ ID No.1 has the amino acid sequence of SEQ ID No.10.
3. Although claims 12 and 13 have been drafted as separate independent claims, they both relate to a recombinant DNA construct comprising a nucleic acid

molecule according to claims 5 and 6. They differ from each other only in that in claim 13 said molecule is inserted in the antisens orientation. Thus, it appears appropriate to amend said claims in terms of one single independent claim and followed by a dependent claim (Rule 6.4 PCT).

4. Claim 18 refers to a nucleic acid molecule according to claims 1 to 8. It seems appropriate to move this claim closer to claims 1 to 8.
5. Claim 20 refers to a polypeptide according to claims 10 or 11. It seems appropriate to move this claim closer to claims 10 and 11.
6. Figure 31 contains text matter (Rule 11.11a).

VIII. Certain observations on the international application

1. The expressions "...comprises..." or "...comprising..." in claims 1-3, 6, 33 and 34 do not clearly define the scope of the claims. Thus, the expressions "comprises" and "comprising" should be replaced with "consists of" and "consisting of", respectively (Article 6 PCT).
3. The terms "fragments" and "derivatives" in claim 3 are vague and not clear, they do not refer to any technical feature, and can be subject to interpretation. A derivative of a known sequence can diverge from said sequence in such a way that the resulting product completely differs from the original one (Art. 6 PCT).
4. Claim 8 refers to an antisens molecule comprising a nucleic acid molecule. The term "antisens molecule" is vague and not clear, and can be subject to interpretation. Moreover, claim 8 attempts to define said subject-matter in terms of a result to be achieved ("capable of hybridising..."). Such a definition is only allowable in case the invention can only be defined in such terms. However, this prerequisite is not met by the instant case since a nucleic acid is a chemical compound which has to be characterised by structural features. Therefore, claim 7 do not meet the requirements of Article 6 PCT (see also Guidelines C-III, 4.7 PCT). Moreover, it should be noted that under low stringency conditions any nucleic acid molecule can hybridise to a specific sequence.

5. The subject-matter of claims 23 and 24 refers to a *C. albicans* cell comprising an induced mutation in the DNA sequence encoding the polypeptide according to claim 10. However, the specification of the present application does not mention any induced mutation in the DNA of *C. albicans* cells. Thus, the subject-matter of claims 23 and 24 is not based on the description (Art. 6-support PCT).

Moreover the present application does not give any indication how to induce mutations in a specific DNA sequence of *C. albicans* cells, which mutation results in overexpression or underexpression of a specific polypeptide. Thus, the man skilled in the art, faced with the problem of providing *C. albicans* cells comprising such an induced mutation would not be able to perform it without the need of applying intensive experimentations of undue burden. Therefore, with respect to claims 23 and 24, the present application does not meet the requirements of Article 5 PCT.

6. The subject-matter of claim 32 refers to an antibody capable of binding to a polypeptide according to claim 10 or 11. Said subject-matter is not clearly defined in that it cannot be ruled out that known antibodies fall within the scope of said claim.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/31, C07K 14/40, G01N 33/50, C12Q 1/68, A61K 31/7088, 38/16, C07K 16/14		A2	(11) International Publication Number: WO 00/34481 (43) International Publication Date: 15 June 2000 (15.06.00)
(21) International Application Number: PCT/EP99/09833 (22) International Filing Date: 6 December 1999 (06.12.99) (30) Priority Data: 98204122.0 4 December 1998 (04.12.98) EP (71) Applicant (for all designated States except US): JANSSEN PHARMACEUTICA N.V. [BE/BE]; Turnhoutseweg 30, B-2340 Beerse (BE). (72) Inventors; and (75) Inventors/Applicants (for US only): <u>CONTRERAS, Roland,</u> Henri [BE/BE]; University of Gent, K.L. Ledeganckstraat 35, B-9000 Gent (BE). <u>NELISSEN, Bart</u> [BE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B-2340 Beerse (BE). <u>DE BACKER, Marianne, Denise</u> [BE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B-2340 Beerse (BE). <u>LUYTEN, Walter, Herman,</u> Maria, Louis [BE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B-2340 beerse (BE). <u>VIAENE, Jasmine, Elza</u> [BE/BE]; University of Gent, K.L. Ledeganck Straat 35, B-9000 Gent (BE). <u>LOGGHE, Marc, George</u> [BE/BE]; University of Gent, K.L. Ledeganckstraat 35, B-9000 Gent (BE). <u>VIALARD,</u>		(74) Agent: BOULT WADE TENNANT; 27 Fumival Street, London EC4A 1PQ (GB). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published Without international search report and to be republished upon receipt of that report.	
(54) Title: DRUG TARGETS IN <i>CANDIDA ALBICANS</i>			
(57) Abstract The present invention is concerned with the identification of genes or functional fragments thereof from <i>Candida albicans</i> which are critical for growth and cell division and which genes may be used as selective drug targets to treat <i>Candida albicans</i> associated infections. Novel nucleic acid sequences from <i>Candida albicans</i> are also provided and which encode the polypeptides which are critical for growth of <i>Candida albicans</i> . Methods for the identification of anti-fungal compounds which inhibit fungal or yeast growth are also contemplated.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

- 1 -

DRUG TARGETS IN *CANDIDA ALBICANS*

The present invention is concerned with the identification of genes or functional fragments thereof from *Candida albicans* which are critical for growth and cell division and which genes may be used as selective drug targets to treat *Candida albicans* associated infections. Novel nucleic acid sequences from *Candida albicans* are also provided and which encode the polypeptides which are critical for growth of *Candida albicans*.

Opportunistic infections in immunocompromised hosts represent an increasingly common cause of mortality and morbidity. *Candida* species are among the most commonly identified fungal pathogens associated with such opportunistic infections, with *Candida albicans* being the most common species. Such fungal infections are thus problematical in, for example, AIDS populations in addition to normal healthy women where *Candida albicans* yeasts represent the most common cause of vulvovaginitis.

Although compounds do exist for treating such disorders, such as, amphotericin, these drugs are generally limited in their treatment because of their toxicity and side effects. Therefore, there exists a need for new compounds which may be used to treat *Candida* associated infections in addition to compounds which are selective in their action against *Candida albicans*.

Classical approaches for identifying anti-fungal compounds have relied almost exclusively on inhibition of fungal or yeast growth as an endpoint. Libraries of natural products, semi-synthetic, or synthetic chemicals are screened for their ability to kill or arrest growth of the target pathogen or a related nonpathogenic model organism. These tests are

- 2 -

cumbersome and provide no information about a compound's mechanism of action. The promising lead compounds that emerge from such screens must then be tested for possible host-toxicity and detailed
5 mechanism of action studies must subsequently be conducted to identify the affected molecular target.

The present inventors have now identified a range of nucleic acid sequences from *Candida albicans* which encode polypeptides which are critical for its
10 survival and growth. These sequences represent novel targets which can be incorporated into an assay to selectively identify compounds capable of inhibiting expression of such polypeptides and their potential use in alleviating diseases or conditions associated
15 with *Candida albicans* infection.

Therefore, according to a first aspect of the invention there is provided a nucleic acid molecule encoding a polypeptide which is critical for survival and growth of the yeast *Candida albicans* and which
20 nucleic acid molecule comprises any of the sequences of nucleotides illustrated in any of Sequence ID Nos. 1 to 9.

Whilst the molecules defined herein have been established as being critical for growth and
25 metabolism of *Candida albicans*, for some of the molecules no apparent functionality has been assigned by virtue of the fact that no functionally related sequences in other prokaryotic or eukaryotic organism can be found in respective databases. Thus,
30 advantageously these sequences may be species specific in which case they may be used as selective targets for treatment of diseases mediated by *Candida Albicans* infection. Thus, in one aspect of the invention the nucleic acid molecules preferably
35 comprise the sequences identified in sequence ID Nos. 1, 4, 5 to 9.

- 3 -

In another aspect of the invention the sequences have been arranged functionally and of nucleotides illustrated in Sequence ID Nos. 2 or 3 are preferred and even more preferably in Sequence ID No. 2 and
5 fragments or derivatives of said nucleic acid molecules.

Letters utilised in the sequences according to the invention which are not recognisable as letters of the genetic code signify a position in the nucleic
10 acid sequence where one or more of bases A, G, C or T can occupy the nucleotide position. Representative letters used to identify the range of bases which can be used are as follows:

15	M:	A or C
	R:	A or G
	W:	A or T
	S:	C or G
	Y:	C or T
20	K:	G or T
	V:	A or C or G
	H:	A or C or T
	D:	A or G or T
	B:	C or G or T
25	N:	G or A or T or C

In one embodiment of each of the above identified aspects of the invention the nucleic acid may comprise a mRNA molecule or alternatively a DNA and preferably
30 a cDNA molecule.

Also provided by the present invention is a nucleic acid molecule capable of hybridising to the nucleic acid molecules illustrated in any of Figures 1 to 9 under high stringency conditions such as
35 antisense molecule and which conditions are generally known to those of skill in the art.

Stringency of hybridisation as used herein refers

- 4 -

to conditions under which polynucleic acids are stable. The stability of hybrids is reflected in the melting temperature (T_m) of the hybrids. T_m can be approximated by the formula:

5

$$81.5^{\circ}\text{C} + 16.6(\text{Log}_{10}[\text{Na}^+] + 0.41 (\% \text{G\&C}) - 600\text{L/L}$$

wherein L is the length of the hybrids in nucleotides. T_m decreases approximately by 1-1.5°C with every 1% decrease in sequence homology.

10

The term "stringency" refers to the hybridisation conditions wherein a single-stranded nucleic acid joins with a complementary strand when the purine or pyrimidine bases therein pair with their corresponding base by hydrogen bonding. High stringency conditions favour homologous base pairing whereas low stringency conditions disfavour non-homologous base pairing.

15

"Low stringency" conditions comprise, for example, a temperature of about 37°C or less, a formamide concentration of less than about 50%, and a moderate to low salt (SSC) concentration; or, alternatively, a temperature of about 50°C or less, and a moderate to high salt (SSPE) concentration, for example 1M NaCl.

20

"High stringency" conditions comprise, for example, a temperature of about 42°C or less, a formamide concentration of less than about 20%, and a low salt (SSC) concentration; or, alternatively, a temperature of about 65°C, or less, and a low salt (SSPE) concentration. For example, high stringency conditions comprise hybridization in 0.5 M NaH_2PO_4 , 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C (Ausubel, F.M. et al. Current Protocols in Molecular Biology, Vol. I, 1989; Green Inc. New York, at 2.10.3).

30

35

"SSC" comprises a hybridization and wash solution. A stock 20X SSC solution contains 3M sodium

- 5 -

chloride, 0.3M sodium citrate, pH 7.0.

"SSPE" comprises a hybridization and wash solution. A 1X SSPE solution contains 180 mM NaCl, 9mM Na₂HPO₄ and 1 mM EDTA, pH 7.4.

5 The nucleic acid capable of hybridising to nucleic acid molecules according to the invention will generally be at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the nucleotide sequences illustrated in any of Figures 1
10 to 9.

 The DNA molecules according to the invention may, advantageously, be included in a suitable expression vector to express polypeptides encoded therefrom in a suitable host which are critical for growth and
15 survival of *Candida albicans*.

 An expression vector according to the invention includes a vector having a nucleic acid according to the invention operably linked to regulatory sequences, such as promoter regions, that are capable of
20 effecting expression of said DNA fragments. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. Such vectors may be transformed into a suitable host
25 cell to provide for expression of a polypeptide according to the invention. Thus, in a further aspect, the invention provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell, transformed,
30 transfected or infected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and recovering the expressed polypeptides.

35 The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression

- 6 -

of said nucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable markers, such as, for example, ampicillin resistance.

5 Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and
10 for translation initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon
15 AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

20 Polynucleotides according to the invention may be inserted into the vectors described in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense nucleic acids may be produced by synthetic means.

25 In accordance with the present invention, a defined nucleic acid includes not only the identical nucleic acid but also any minor base variations including in particular, substitutions in bases which result in a synonymous codon (a different codon specifying the same amino acid residue) due to the
30 degenerate code. The term "nucleic acid sequence" also includes the complementary sequence to any single stranded sequence given regarding base variations.

35 The present invention also comprises within its scope proteins or polypeptides expressed by the nucleic acid molecules according to the invention or a functional equivalent, derivative or bioprecursor thereof.

- 7 -

The present invention also advantageously provides nucleic acid sequences of at least approximately 10 contiguous nucleotides of a nucleic acid according to the invention and preferably from 10 to approximately 120 nucleotides. In another aspect of the invention, nucleotide acid sequences are provided from 10 to 50 nucleotides. These sequences may, advantageously be used as probes or primers to initiate replication, or the like. Such nucleic acid sequences may be produced according to techniques well known in the art, such as by recombinant or synthetic means. They may also be used in diagnostic kits or the like for detecting the presence of a nucleic acid according to the invention. These tests generally comprise contacting the probe with the sample under hybridising conditions and detecting for the presence of any duplex or triplex formation between the probe and any nucleic acid in the sample.

According to the present invention, these probes may be anchored to a solid support. Preferably, they are present on an array so that multiple probes can simultaneously hybridize to a single biological sample. The probes can be spotted onto the array or synthesized *in situ* on the array. See Lockhart et al., Nature Biotechnology, Vol. 14, December 1996, "Expression monitoring by hybridization to high-density oligonucleotide arrays." A single array can contain more than up to more than a million different probes in discrete locations.

Advantageously, the nucleic acid sequences, according to the invention may be produced using such recombinant or synthetic means, such as for example using PCR cloning mechanisms which generally involve making a pair of primers, which may be between approximately 10 to 120 nucleotides to a region of the gene which is desired to be cloned, bringing the primers into contact with mRNA, cDNA, or genomic DNA

- 8 -

from a cell, performing a polymerase chain reaction under conditions which bring about amplification of the desired region, isolated the amplified region or fragment and recovering the amplified DNA. Generally, such techniques as defined herein are well known in the art, such as described in Sambrook et al (Molecular Cloning: a Laboratory Manual, 1989).

The nucleic acids or oligonucleotides according to the invention may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S , enzyme labels or other protein labels such as biotin or fluorescent markers. Such labels may be added to the nucleic acids or oligonucleotides of the invention and may be detected using known techniques *per se*.

The polypeptide or protein according to the invention includes all possible amino acid variants encoded by the nucleic acid molecule according to the invention including a polypeptide encoded by said molecule and having conservative amino acid changes. Polypeptides according to the invention further include variants of such sequences, including naturally occurring allelic variants which are substantially homologous to said polypeptides. In this context, substantial homology is regarded as a sequence which has at least 70%, preferably 80 or 90% amino acid homology with the polypeptides encoded by the nucleic acid molecules according to the invention.

Nucleic acids and polypeptides which are particularly preferred are those comprising the sequences of nucleotides illustrated in figures 1 to 3 and polypeptides illustrated in figures 14 to 16. However, a particularly preferred nucleic acid comprises the sequences of nucleotides illustrated in Figures 2 and/or 3, and their corresponding amino acid sequences identified in Figures 15 and 16.

Nucleotide sequences according to the invention

- 9 -

are particularly advantageous as selective therapeutic targets for treating *Candida albicans* associated infections. For example, an antisense nucleic acid capable of binding to the nucleic acid sequence
5 illustrated in any of Figures 1 to 9 may be used to selectively inhibit expression of the corresponding polypeptides, leading to impaired growth of the *Candida albicans* with reductions of associated illnesses or diseases.

10 The nucleic acid molecule or the polypeptide according to the invention may be used as a medicament, or in the preparation of a medicament, for treating diseases or conditions associated with *Candida albicans* infection.

15 Advantageously, the nucleic acid molecule or the polypeptide according to the invention may be provided in a pharmaceutical composition together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

20 The present invention is further directed to inhibiting expression of nucleic acids according to the invention *in vivo* by the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation
25 of antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion or the mature protein sequence, which encodes for the protein of the present invention, is used to design an antisense RNA
30 oligonucleotide of from 10 to 50 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple-helix - see Lee et al. Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241:456 (1988);
35 and Dervan et al., Science, 251: 1360 (1991), thereby preventing transcription and the production of the

- 10 -

corresponding protein. The antisense RNA oligonucleotide hybridises to the mRNA *in vivo* and blocks translation of an mRNA molecule into the corresponding protein (antisense - Okano, J. Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)).

Antibodies to the protein or polypeptide of the present invention may, advantageously, be prepared by techniques which are known in the art. For example, polyclonal antibodies may be prepared by inoculating a host animal, such as a mouse, with the polypeptide according to the invention or an epitope thereof and recovering immune serum. Monoclonal antibodies may be prepared according to known techniques such as described by Kohler R. and Milstein C., Nature (1975) 256, 495-497.

Antibodies according to the invention may also be used in a method of detecting for the presence of a polypeptide according to the invention, which method comprises reacting the antibody with a sample and identifying any protein bound to said antibody. A kit may also be provided for performing said method which comprises an antibody according to the invention and means for reacting the antibody with said sample.

Proteins which interact with the polypeptide of the invention may be identified by investigating protein-protein interactions using the two-hybrid vector system first proposed by Chien *et al.* (1991).

This technique is based on functional reconstitution *in vivo* of a transcription factor which activates a reporter gene. More particularly the technique comprises providing an appropriate host cell with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain, expressing in the host cell a first hybrid DNA

- 11 -

sequence encoding a first fusion of a fragment or all
of a nucleic acid sequence according to the invention
and either said DNA binding domain or said activating
domain of the transcription factor, expressing in the
5 host at least one second hybrid DNA sequence, such as
a library or the like, encoding putative binding
proteins to be investigated together with the DNA
binding or activating domain of the transcription
factor which is not incorporated in the first fusion;
10 detecting any binding of the proteins to be
investigated with a protein according to the invention
by detecting for the presence of any reporter gene
product in the host cell; optionally isolating second
hybrid DNA sequences encoding the binding protein.

15 An example of such a technique utilises the GAL4
protein in yeast. GAL4 is a transcriptional activator
of galactose metabolism in yeast and has a separate
domain for binding to activators upstream of the
galactose metabolising genes as well as a protein
20 binding domain. Nucleotide vectors may be
constructed, one of which comprises the nucleotide
residues encoding the DNA binding domain of GAL4.
These binding domain residues may be fused to a known
protein encoding sequence, such as for example the
25 nucleic acids according to the invention. The other
vector comprises the residues encoding the protein
binding domain of GAL4. These residues are fused to
residues encoding a test protein. Any interaction
between polypeptides encoded by the nucleic acid
30 according to the invention and the protein to be
tested leads to transcriptional activation of a
reporter molecule in a GAL4 transcription deficient
yeast cell into which the vectors have been
transformed. Preferably, a reporter molecule such as
35 β -galactosidase is activated upon restoration of
transcription of the yeast galactose metabolism genes.

Further provided by the present invention is one

- 12 -

or more *Candida albicans* cells comprising an induced mutation in the DNA sequence encoding the polypeptide according to the invention.

5 A further aspect of the invention provides a method of identifying compounds which selectively inhibit or interfere with the expression, the functionality of polypeptides expressed from the nucleotides sequences illustrated in any of Figures 1 to 9 or the metabolic pathways in which these
10 polypeptides are involved and which are critical for growth and survival of *Candida albicans*, which method comprises (a) contacting a compound to be tested with one or more *Candida albicans* cells having a mutation in a nucleic acid molecule according to the invention
15 which mutation results in overexpression or underexpression of said polypeptides in addition to one or more wild type *Candida* cells, (b) monitoring the growth and/or activity of said mutated cell compared to said wild type wherein differential growth
20 or activity of said one or more mutated *Candida* cells provides an indication of selective action of said compound on said polypeptide or another polypeptide in the same or a parallel pathway.

Compounds identifiable or identified using the
25 method according to the invention, may advantageously be used as a medicament, or in the preparation of a medicament to treat diseases or conditions associated with *Candida albicans* infection. These compounds may also advantageously be included in a pharmaceutical
30 composition together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

A further aspect of the invention provides a method of identifying DNA sequences from a cell or organism which DNA encodes polypeptides which are
35 critical for growth or survival, which method comprises (a) preparing a cDNA or genomic library from

- 13 -

said cell or organism in a suitable expression vector which vector is such that it can either integrate into the genome in said cell or that it permits transcription of antisense RNA from the nucleotide sequences in said cDNA or genomic library, (b) selecting transformants exhibiting impaired growth and determining the nucleotide sequence of the cDNA or genomic sequence from the library included in the vector from said transformant. Preferably, the cell or organism may be any yeast or filamentous fungus, such as, for example, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* or *Candida albicans*.

A further aspect of the invention provides a pharmaceutical composition comprising any of a compound, an antisense molecule or an antibody according to the invention together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

The antisense molecules or indeed the compounds identified as agonists or antagonists of the nucleic acids or polypeptides according to the invention may be used in the form of a pharmaceutical composition, which may be prepared according to procedures well known in the art. Preferred compositions include a pharmaceutically acceptable vehicle or diluent or excipient, such as for example, a physiological saline solution. Other pharmaceutically acceptable carriers including other non-toxic salts, sterile water or the like may also be used. A suitable buffer may also be present allowing the compositions to be lyophilized and stored in sterile conditions prior to reconstitution by the addition of sterile water for subsequent administration. Incorporation of the polypeptides of the invention into a solid or semi-solid biologically compatible matrix may be carried out which can be implanted into tissues requiring treatment.

- 14 -

The carrier can also contain other pharmaceutically acceptable excipients for modifying other conditions such as pH, osmolarity, viscosity, sterility, lipophilicity, solubility or the like.

5 Pharmaceutically acceptable excipients which permit sustained or delayed release following administration may also be included.

The polypeptides, the nucleic acid molecules or compounds according to the invention may be
10 administered orally. In this embodiment they may be encapsulated and combined with suitable carriers in solid dosage forms which would be well known to those skilled in the art.

As would be well known to those of skill in the
15 art, the specific dosage regime may be calculated according to the body surface area of the patient or the volume of body space to be occupied, dependent upon the particular route of administration to be used. The amount of the composition actually
20 administered will, however, be determined by a medical practitioner, based on the circumstances pertaining to the disorder to be treated, such as the severity of the symptoms, the composition to be administered, the age, weight, and response of the individual patient
25 and the chosen route of administration.

The present invention may be more clearly understood with reference to the accompanying example, which is purely exemplary, with reference to the accompanying drawings, wherein

30

Figures 1 and 2: are nucleotide sequences isolated from *Candida albicans* and which have an identified function based on sequence homology with proteins from other organisms and which sequences are not present in the public domain.

35

- 15 -

- Figures 3 : is a nucleotide sequence isolated from *Candida albicans* and which has an identified function based on sequence homology with proteins from other organisms and which sequence is partially present in the public domain.
- Figures 4 : is a nucleotide sequence of previously unknown function isolated from *Candida albicans* and which is partially present in the public domain.
- Figures 5 to 9 : are nucleotide sequences of previously unknown function isolated from *Candida albicans*.
- Figure 10 : is a diagrammatic representation of plasmid pGAL1PNiST-1.
- Figure 11 : is a nucleotide sequence of plasmid pGAL1PNiST-1 of Figure 10.
- Figure 12 : is a diagrammatic representation of plasmid pGAL1PSiST-1.
- Figure 13 : is a nucleotide sequence of plasmid pGAL1PSiST-1 of Figure 12.
- Figures 14 to 20: are amino acid sequences of the appropriately corresponding DNA sequences illustrated in Figures 1 to 9 with reference to Table 1.
- Figures 21 to 27: are growth curves of *Candida albicans* strains showing antisense

- 16 -

induced reduction in growth.

Figures 28 to 31: are growth curves of *Candida albicans* strains including knock-outs in the relevant gene identified.

Example 1

Identification of novel drug targets in *C. albicans* by anti-sense and disruptive integration

The principle of the approach is based on the fact that when a particular *C. albicans* mRNA is inhibited by producing the complementary anti-sense RNA, the corresponding protein will decrease. If this protein is critical for growth or survival, the cell producing the anti-sense RNA will grow more slowly or will die.

Since anti-sense inhibition occurs at mRNA level, the gene copy number is irrelevant, thus allowing applications of the strategy even in diploid organisms.

Anti-sense RNA is endogenously produced from an integrative or episomal plasmid with an inducible promoter; induction of the promoter leads to the production of an RNA encoded by the insert of the plasmid. This insert will differ from one plasmid to another in the library. The inserts will be derived from genomic DNA fragments or from cDNA to cover to the extent possible- the entire genome.

The vector is a proprietary vector allowing integration by homologous recombination at either the homologous insert or promoter sequence in the *Candida* genome. After introducing plasmids from cDNA or genomic libraries into *C. albicans*, transformants are screened for impaired growth after promoter (& thus anti-sense) induction in the presence of lithium

- 17 -

acetate. Lithium acetate prolongs the G1 phase and thus allows anti-sense to act during a prolonged period of time during the cell cycle. Transformants which show impaired growth in both induced and non-induced media, thus showing a growth defect due to integrative disruption, are selected as well.

Transformants showing impaired growth are supposed to contain plasmids which produce anti-sense RNA to mRNAs critical for growth or survival. Growth is monitored by measuring growth-curves over a period of time in a device (Bioscreen Analyzer, Labsystems) which allows simultaneous measurement of growth-curves of 200 transformants.

Subsequently plasmids can be recovered from the transformants and the sequence of their inserts determined, thus revealing which mRNA they inhibit. In order to be able to recover the genomic or cDNA insert which has integrated into the *Candida* genome, genomic DNA is isolated, cut with an enzyme which cuts only once into the library vector (and estimated approx. every 4096 bp in the genome) and religated. PCR with primers flanking the insert will yield (partial) genomic or cDNA inserts as PCR fragments which can directly be sequenced. This PCR analysis (on ligation reaction) will also show us how many integrations occurred. Alternatively the ligation reaction is transformed to *E. coli* and PCR analysis is performed on colonies or on plasmid DNA derived thereof.

This method is employed for a genome-wide search for novel *C. albicans* genes which are important for growth or survival.

Materials & Methods

Construction of pGAL1PNiST-1

The backbone of the pGAL1PNiST-1 vector (integrative anti-sense *SfiI*-*NotI* vector) is

- 18 -

pGEM11Zf(+) (Promega Inc.). First, the CaMAL2
EcoRI/SalI promoter fragment from pDBV50 (D.H. Brown
et al. 1996) was ligated into EcoRI/SalI-opened
pGEM11Zf(+) resulting in the intermediate construct
5 pGEMMAL2P-1. Into the latter (MscI/CIP) the CaURA3
selection marker was cloned as a Eco47III/XmnI
fragment derived from pRM2. The resulting pGEMMAL2P-2
vector was NotI/HindIII opened in order to accept the
NotI-stuffer-SfiI cassette from pPCK1NiSCYCT-1
10 (EagI/HindIII fragment): pMAL2PNiST-1. Finally, the
plasmid pGAL1PNiST-1 was constructed by exchanging the
SalI/Ecl136II MAL2 promoter in pMAL2PNiST-1 by the
XhoI/SmaI GAL1 promoter fragment derived from
pRM2GAL1P.

15

Construction of pGAL1PSiST-1

The vector pGAL1PSiST-1 was created for cloning
the small genomic DNA fragments (flanked by SfiI
sites) behind the GAL1 promoter. The only difference
20 with pGAL1PNiST-1 is that the hIFN β (stuffer fragment)
insert fragment in pGAL1PSiST-1 is flanked by two SfiI
sites instead of a SfiI and a NotI site as in
pGAL1PNiST-1. To construct pGAL1PSiST-1 the EcoRI-
HindIII fragment, containing hIFN β flanked by a SfiI
25 and a NotI site, of pMAL2pHiET-3 (unpublished) was
exchanged by the EcoRI-HindIII fragment, containing
hIFN β flanked by two SfiI sites, from YCp50S-S (an *E.*
coli / *S. cerevisiae* shuttle vector derived from the
plasmid YCp50, which is deposited in the ATCC
30 collection (number 37419; Thrash et al., 1985); an
EcoRI-HindIII fragment, containing the gene hIFN β ,
which is flanked by two SfiI sites, was inserted in
YCp50, creating YCp50S-S), resulting into plasmid
pMAL2PSiST-1. The MAL2 promoter from pMAL2PSiST-1 (by
35 a NaeI-balI digest) was further replaced by the GAL1
promoter from pGAL1PNiST-1 (via a XhoI-FSPI digest),

- 19 -

creating the vector pGAL1PSiST-1.

***Candida albicans* genomic library**

**** Preparation of the genomic DNA fragments***

5 A *Candida albicans* genomic DNA library with small DNA
fragments (400 to 1,000 bp) was prepared. Genomic DNA
of *Candida albicans* B2630 was isolated following a
modified protocol of Blin and Stafford (1976). The
10 quality of the isolated genomic DNA was checked by gel
electrophoresis. Undigested DNA was located on the gel
above the marker band of 26,282 bp. A little smear,
caused by fragmentation of the DNA, was present.
To obtain enrichment for genomic DNA fragments of the
desired size, the genomic DNA was partially digested.
15 Several restriction enzymes (*AluI*, *HaeIII* and *RsaI*;
all creating blunt ends) were tried out. The
appropriate digest conditions have been determined by
titration of the enzyme. Enrichment of small DNA
fragments was obtained with 70 units of *AluI* on 10 µg
20 of genomic DNA for 20 min. T4 DNA polymerase
(Boehringer) and dNTPs (Boehringer) were added to
polish the DNA ends. After extraction with phenol-
chloroform the digest was size-fractionated on an
agarose gel. The genomic DNA fragments with a length
25 of 500 to 1,250 bp were eluted from the gel by
centrifugal filtration (Zhu et al., 1985). *SfiI*
adaptors (5' GTTGGCCTTTT) or (5' AGGCCAAC) were
attached to the DNA ends (blunt) to facilitate cloning
of the fragments into the vector. Therefore, a 8-mer
30 and 11-mer oligonucleotide (comprising the *SfiI* site)
were kinased and annealed. After ligation of these
adaptors to the DNA fragments a second size-
fractionation was performed on an agarose gel. The
DNA fragments of 400 to 1150 bp were eluted from the
35 gel by centrifugal filtration.

**** Preparation of the pGAL1PSiST-1 vector fragment***

- 20 -

The small genomic DNA fragments were cloned after the GAL1 promoter in the vector pGAL1PSiST-1. Qiagen-purified pGAL1PSiST-1 plasmid DNA was digested with *SfiI* and the largest vector fragment eluted from the gel by centrifugal filtration (Zhu et al., 1985).
5 Ligation with a control DNA fragment, flanked by *SfiI* sites, was performed as a control. The ligation mix was electroporated to MC1061 *E. coli* cells. Plasmid DNA of 24 clones was analyzed. In all cases the
10 control fragment was inserted in the pGAL1PSiST-1 vector fragment.

** Upscaling*

All genomic DNA fragments (450 ng) were ligated into the pGAL1PSiST-1 vector (20 ng). After
15 electroporation at 2500V, 40µF circa 400,000 clones were obtained. These clones were pooled into three groups and stored as glycerol slants. Also Qiagen-purified DNA was prepared from these clones. A clone analysis showed an average insert length of 600 bp and
20 a percentage of 91 for clones with an insert. The size of the library corresponds to 5 times the diploid genome. The genomic DNA inserts are sense or anti-sense orientated in the vector.

25 ***Candida albicans* cDNA library**

Total RNA was extracted from *Candida albicans* B2630 grown on respectively minimal (SD) and rich (YPD) medium as described by Chirgwin et al. in Sambrook et al 1996. mRNA was prepared from total RNA
30 using the Invitrogen Fast Track procedure.

First strand cDNA is synthesised with the Superscript Reverse Transcriptase (BRL) and with an oligo dT-NotI Primer adapter. After second strand synthesis, cDNA is polished with Klenow enzyme and
35 purified over a Sephacryl S-400 spun column. Phosphorylated *SfiI* adapters are then ligated to the

- 21 -

cDNA, followed by digestion with the *NotI* restriction enzyme. The *SfiI/NotI* cDNA is then purified and sized on a Biogel column A150M.

First fraction contains approximately 38,720 clones by transformation, the second fraction only 1540 clones. Clone analysis:
Fr. I: 22/24 inserts, 16 ³ 1000 bp, 4 ³ 2000 bp, average size: 1500 bp.
Fr. II: 9/12 inserts, 3 ³ 1000 bp, average size: 960 bp cDNA was ligated in a *NotI/SfiI* opened pGAL1PNiST-1 vector (anti-sense)

***Candida* transformation**

The host strain used for transformation is a *C. albicans* *ura3* mutant, CAI-4, which contains a deletion in orotidine-5'-phosphate decarboxylase and was obtained from William Fonzi, Georgetown University (Fonzi and Irwin). CAI-4 was transformed with the above described cDNA library or genomic library using the *Pichia* spheroplast module (Invitrogen). Resulting transformants were plated on minimal medium supplemented with glucose (SD, 0.67% or 1.34% Yeast Nitrogen base w/o amino acids + 2% glucose) plates and incubated for 2-3 days at 30°C.

Screening for mutants

Starter cultures were set up by inoculating each colony in 1 ml SD medium and incubating overnight at 30°C and 300 rpm. Cell densities were determined using a Coulter counter (Coulter Z1; Coulter electronics limited). 250.000 cells/ml were inoculated in 1 ml SD medium and cultures were incubated for 24 hours at 30°C and 300 rpm. Cultures were washed in minimal medium without glucose (S) and the pellet resuspended in 650 µl S medium. 8 µl of this culture is used for inoculating 400 µl cultures in a Honeywell-100 plate

- 22 -

(Bioscreen analyzer; Labsystems). Each transformant was grown during three days in S medium containing LiAc; pH 6.0, with 2% glucose/2% maltose or 2% galactose/2% maltose respectively while shaking every 3 minutes for 20 seconds. Optical densities were measured every hour during three consecutive days and growth curves were generated (Bioscreen analyzer; Labsystems).

Growth curves of transformants grown in respectively anti-sense non-inducing (glucose/maltose) and inducing (galactose/maltose) medium are compared and those transformants showing impaired growth upon anti-sense induction are selected for further analysis. Transformants showing impaired growth by virtue of integration into a critical gene are also selected.

Isolation of genomic or cDNA inserts

Putatively interesting transformants are grown in 1.5 ml SD overnight and genomic DNA is isolated using the Nucleon MI Yeast kit (Clontech). Concentration of genomic DNA is estimated by analyzing a sample on an agarose gel.

20 ng of genomic DNA is digested for three hours with an enzyme that cuts uniquely in the library vector (SacI for the genomic library; PstI for the cDNA library) and treated with RNase. Samples are phenol/chloroform extracted and precipitated using NaOAc/ethanol.

The resulting pellet is resuspended in 500 μ l ligation mixture (1 x ligation buffer and 4 units of T4 DNA ligase; both from Boehringer) and incubated overnight at 16°C.

After denaturation (20 min 65°C), purification (phenol/chloroform extraction) and precipitation (NaOAc/ethanol) the pellet is resuspended in 10 μ l MilliQ (Millipore) water.

- 23 -

PCR analysis

Inverse PCR is performed on 1 μ l of the precipitated ligation reaction using library vector specific primers (oligo23 5' TGC-AGC-TCG-ACC-TCG-ACT-G 3' and oligo25 5' GCG-TGA-ATG-TAA-GCG-TGA-C 3' for the genomic library; 3pGALNistPCR primer :5'TGAGCAGCTCGCCGTCGCGC 3' and 5pGALNistPCR primer: 5'GAGTTATACCCTGCAGCTCGAC 3' for the cDNA library; both from Eurogentec) for 30 cycles each consisting of (a) 1 min at 95 °C, (b) 1 min at 57 °C, and (c) 3 min at 72 °C. In the reaction mixture 2.5 units of Taq polymerase (Boehringer) with TaqStart antibody (Clontech) (1:1) were used, and the final concentrations were 0.2 μ M of each primer, 3 mM MgCl₂ (Perkin Elmer Cetus) and 200 μ M dNTPs (Perkin Elmer Cetus). PCR was performed in a Robocycler (Stratagene).

Sequence determination

Resulting PCR products were purified using PCR purification kit (Qiagen) and were quantified by comparison of band intensity on EtBr stained agarose gel with the intensity of DNA marker bands. The amount of PCR product (expressed in ng) used in the sequencing reaction is calculated as the length of the PCR product in basepairs divided by 10. Sequencing reactions were performed using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit according to the instructions of the manufacturer (PE Applied Biosystems, Foster City, CA) except for the following modifications.

The total reaction volume was reduced to 15 μ l. Reaction volume of individual reagents were changed accordingly. 6.0 μ l Terminator Ready Reaction Mix was replaced by a mixture of 3.0 μ l Terminator Ready Reaction Mix + 3.0 μ l Half Term (GENPAK Limited,

- 24 -

Brighton, UK). After cycle sequencing, reaction mixtures were purified over Sephadex G50 columns prepared on Multiscreen HV opaque microtiter plates (Millipore, Molsheim, Fr) and were dried in a speedVac. Reaction products were resuspended in 3 μ l loading buffer. Following denaturation for 2 min at 95°C, 1 μ l of sample was applied on a 5% Long Ranger Gel (36 cm well-to-read) prepared from Singel Packs according to the supplier's instructions (FMC BioProducts, Rockland, ME). Samples were run for 7 hours 2X run on a ABI 377XL DNA sequencer. Data collection version 2.0 and Sequence analysis version 3.0 (for basecalling) software packages are from PE Applied Biosystems. Resulting sequence text files were copied onto a server for further analysis.

Sequence analysis

Nucleotide sequences were imported in the VectorNTI software package (InforMax Inc, North Bethesda, MD, USA), and the vector and insert regions of the sequences were identified. Sequence similarity searches against public and commercial sequence databases were performed with the BLAST software package (Altschul et al., 1990) version 1.4. Both the original nucleotide sequence and the six-frame conceptual translations of the insert region were used as query sequences. The used public databases were the EMBL nucleotide sequence database (Stoesser et al., 1998), the SWISS-PROT protein sequence database and its supplement TrEMBL (Bairoch and Apweiler, 1998), and the ALCES *Candida albicans* sequence database (Stanford University, University of Minnesota). The commercial sequence databases used were the LifeSeq® human and PathoSeq® microbial genomic databases (Incyte Pharmaceuticals Inc., Palo Alto, CA, USA), and the GENESEQ patent sequence database (Derwent, London,

- 25 -

UK). Three major results were obtained on the basis of the sequence similarity searches: function, novelty, and specificity. A putative function was deduced on the basis of the similarity with sequences with a known function, the novelty was based on the absence or presence of the sequences in public databases, and the specificity was based on the similarity with vertebrate homologues.

10 **Methods**

Blastx of the nucleic acid sequences against the appropriate protein databases: Swiss-Prot for clones of which the complete sequence is present in the public domain, and paorfp (PathoSeq™) for clones of which the complete sequences is not present in the public domain.

The protein to which the translated nucleic acid sequence corresponds to is used as a starting point. The differences between this protein and our translated nucleic acid sequences are marked with a double line and annotated above the protein sequence. The following symbols are used:

a one-letter amino acid code or the ambiguity code X is used if our translated nucleic acid sequence has another amino acid on a certain position,

the stop codon sign * is used if our translated nucleic acid sequence has a stop codon on a certain position,

The letters fs (frame shift) are used if a frame shift occurs in our translated nucleic acid sequence, and another reading frame is used,

the words ambiguity or ambiguities are used if a part of our translated nucleic acid sequence is present in the proteins, but not visible in the alignments of the blast results,

The phrase "missing sequence" is used if the translated nucleic acid sequence does not comprise

- 26 -

that part of the protein.

Blastx: compares the six-frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database.

5

Gene Knock-outs

To verify that the growth effect was due to the interference with the identified gene and to support the specificity of the antisense effect, single allele knock-outs were made in the identified genes (Figures 28 to 31) using the URA-blaster method (Fonzi and Irwin 1993).

Screening for compounds modulating expression of polypeptides critical for growth and survival of *C. albicans*

The method proposed is based on observations (Sandbaken et al., 1990; Hinnebusch and Liebman 1991; Ribogene PCT WO 95/11969, 1995) suggesting that underexpression or overexpression of any component of a process (e.g. translation) could lead to altered sensitivity to an inhibitor of a relevant step in that process. Such an inhibitor should be more potent against a cell limited by a deficiency in the macromolecule catalyzing that step and/or less potent against a cell containing an excess of that macromolecule, as compared to the wild type (WT) cell.

Mutant yeast strains, for example, have shown that some steps of translation are sensitive to the stoichiometry of macromolecules involved. (Sandbaken et al. 1996). Such strains are more sensitive to compounds which specifically perturb translation (by acting on a component that participates in translation) but are equally sensitive to compounds with other mechanisms of action.

35

This method thus not only provides a means to

- 27 -

identify whether a test compound perturbs a certain process but also an indication of the site at which it exerts its effect. The component which is present in altered form or amount in a cell whose growth is
5 affected by a test compound is potentially the site of action of the test compound.

The assay to be set up involves measurement of growth of an isogenic strain which has been modified only in a certain specific allele, relative to a wild
10 type (WT) *C. albicans* strain, in the presence of R-compounds. Strains can be ones in which the expression of a specific essential protein is impaired upon induction of anti-sense or strains which carry disruptions in an essential gene. An *in silico*
15 approach to finding novel essential genes in *C. albicans* will be performed. A number of essential genes identified in this way will be disrupted (in one allele) and the resulting strains can be used for comparative growth screening.

20

Assay for High Throughput screening for drugs

35 μ l minimal medium (S medium + 2% galactose + 2% maltose) is transferred in a transparent flat-bottomed 96 well plate using an automated pipetting
25 system (Multidrop, Labsystems). A 96-channel pipettor (Hydra, Robbins Scientific) transfers 2.5 μ l of R-compound at 10^{-3} M in DMSO from a stock plate into the assay plate.

The selected *C. albicans* strains (mutant and
30 parent (CAI-4) strain) are stored as glycerol stocks (15%) at -70°C . The strains are streaked out on selective plates (SD medium) and incubated for two days at 30°C . For the parent strain, CAI-4, the medium is always supplemented with 20 $\mu\text{g/ml}$ uridine. A single
35 colony is scooped up and resuspended in 1 ml minimal medium (S medium + 2% galactose + 2% maltose). Cells

- 28 -

are incubated at 30°C for 8 hours while shaking at 250 rpm. A 10 ml culture is inoculated at 250.000 cells/ml. Cultures are incubated at 30°C for 24 hours while shaking at 250 rpm. Cells are counted in
5 Coulter counter and the final culture (S medium + 2% galactose + 2% maltose) is inoculated at 20.000 to 50.000 cells/ml. Cultures are grown at 30°C while shaking at 250 rpm until a final OD of 0.24 (+/- 0.04) 600nm is reached.

10 200 µl of this yeast suspension is added to all wells of MW96 plates containing R-compounds in a 450 (or 250) µl total volume. MW96 plates are incubated (static) at 30°C for 48 hours.

Optical densities are measured after 48 hours.
15 Test growth is expressed as a percentage of positive control growth for both mutant (x) and wild type (y) strains. The ratio (x/y) of these derived variables is calculated.

- 29 -

References

- 5 Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and
D.J. Lipman, 1990, Basic Local Alignment Search Tool,
J. Mol. Biol. 215: 403-410.
- Thrash C., Bankier A. T., Barrell B. G. and Sternglanz
R. (1985) *Proc. Natl. Acad. Sci. USA* **82**: 4374-4378.
- 10 Blin N. and Stafford D. W. (1976) *Nucleic Acids Res.*
3: 2303-2308.
- Zhu J., Kempenaers W., Van Der Straeten D., Contreras
R., and Fiers W. (1985) *Bio/Technology* **3**: 1014-1016.
- 15 Sambrook J., Fritsch E.F. and Maniatis T. (1989)
Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold
Spring Harbor Laboratory, Cold Spring Harbor, NY.
- 20 Fonzi W. and Irwin H. (1993) *Genetics* **134**:717-728.
Altschul S.F, Gish W., Miller W., Myers E.W., Lipman
D.J. (1990) *J. Mol. Biol.* **215**(3):403-410.
- Bairoch A. and Apweiler R. (1998) *Nucleic Acids Res.*
25 **26**(1):38-42.
- Stoesser G., Moseley M.A., Sleep J., McGowran M.,
Garcia-Pastor M., Sterk P. (1998) *Nucleic Acids Res.*
26(1):8-15.
- 30 Chien et al., (1991) *Proc. Natl. Acad. Sci USA* **88**,
9578-9582.
- Sandbaken M.G. Lupisella J.A., DiDomenico B. and
35 Chakraborty K., *J. Biol. Chem.* **265**:15838-15844, 1990.

- 30 -

Hinnebusch A.G. and Liebman S.W., in : The Molecular Biology of the Yeast *Saccharomyces*, Broach J.R., Pringle J.R. and Jones E.W., eds., CSH Laboratory Press; NY 1991.

5

Patent application RiboGene Inc., PCT WO 95./11969, 1995.

10

Brown, D.H. et al., Molecular General Genetics 251 (1) 75-80, 1996.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

15

Arndt, G.M., D. Atkins, M. Patrikakis, and J.G. Izant 1995. Gene regulation by antisense RNA in the fission yeast *Schizosaccharomyces pombe*. Mol. Gen. Genet. 248: 293-300.

20

Atkins, D., and W.L. Gerlach. 1994. Artificial ribozyme and antisense gene expression in *S. cerevisiae*. Antisense research and development 4:109-117.

25

Atkins, D., G.M. Arndt, and J.G. Izant 1994. Antisense gene expression in yeast. Biol. Chem. Hoppe-Seyler. 375:721-729.

30

Atkins, D., M. Patrikakis, J.G. Izant 1995. The *ade6* gene of the fission yeast as a target for antisense and ribozyme RNA-mediated suppression. Antisense Research & Development. 5(4):295-305.

35

Bairoch, A., and R. Apweiler. 1998. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1998. Nucleic Acids Res. 26: 38-42.

- 31 -

- Baudin, A., O. Ozier-Kalogeropoulos, A. Denouel, F. Lacroute, C. Cullin. 1993. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Research*. 21(14):3329-30.
- 5
Blin, N., and D.W. Stafford. 1976. *Nucleic Acids Res.* 3: 2303-2308.
- Dujon, B. 1998. European Functional Analysis Network (EUROFAN) and the functional analysis of the
10 *Saccharomyces cerevisiae* genome. *Electrophoresis*. 19:617-624.
- Del Rosario, M., J.C. Stephans, J. Zakel, J. Escobedo,
15 and K. Giese. 1996. Positive selection system to screen for inhibitors of human immunodeficiency virus-1 transcription. *Nature Biotechnology*. 14:1592-1596.
- Fairhead, C., A. Thierry, F. Denis, M. Eck, and B.
20 Dujon. 1998. "Mass-murder" of ORFs from three regions of chromosome XI from *Saccharomyces cerevisiae*. *Gene*. 223:33-46.
- Ferbeyre, G., J. Bratty, H. Chen, R. Cedergren. 1996.
25 Cell cycle arrest promotes trans-hammerhead ribozyme action in yeast. *Journal of Biological Chemistry*. 271(32):19318-23.
- Fleischmann, R.D., M.D. Adams, O. White, R.A. Clayton,
30 E.F. Kirkness, A.R. Kerlavage, C.J. Bult, J.F. Tomb, B.A. Dougherty, J.M. Merrick. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae*. *Science*. 269: 496-512
- Fonzi, W.A., and M.Y. Irwin. 1993. Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* 134:717-728.
- 35

- 32 -

Fraser, C.M., J.D. Gocayne, O. White, M.D. Adams, R.A. Clayton, R.D. Fleischmann, C.J. Bult, A.R. Kervalage, G. Sutton, J.M. Kelley. 1995. The minimal gene complement of *Mycoplasma genitalium*. Science 270:397-403.

Giese, K. 1997. Method and construct for screening for inhibitors of transcriptional activation. International patent application WO 97/10360.

Hahn, S. Guarente L. 1988. Yeast HAP2 and HAP3: transcriptional activators in a heteromeric complex. Science. 240(4850):317-21

Heid, C.A., J. Stevens, K.J. Livak, and P.M. Williams. 1996. Real time quantitative PCR. Genome Methods 6:986-994.

Jayaram, M., A. Sutton, J.R. Broach. 1985. Properties of REP3: a cis-acting locus required for stable propagation of the *Saccharomyces cerevisiae* plasmid 2 microns circle. Molecular & Cellular Biology. 5(9):2466-75.

Jeong, S.W., W.H. Lang, R.H. Reeder. 1996. The yeast transcription terminator for RNA polymerase I is designed to prevent polymerase slippage. Journal of Biological Chemistry. 271(27):16104-10.

Johnson, F.B., D.A. Sinclair, and L. Guarente. 1999. Molecular Biology of aging. Cell. 96:291-302.

Leuker, C.E., A. Sonneborn, S. Delbruck, J.F. Ernst. 1997. Sequence and promoter regulation of the PCK1 gene encoding phosphoenolpyruvate carboxykinase of the fungal pathogen *Candida albicans*. Gene. 192(2):235-40.

- 33 -

Lie, Y.S., and C.J. Petropoulos. 1998. Advances in quantitative PCR technology: 5' nuclease assays. *Current Opinion in Biotechnology* 9:43-48.

- 5 Livak, K.J., S.J. Flood, J. Marmaro, W. Giusti, K. Deetz. 1995. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *Genome Research*. 4(6):357-62.
- 10 Mandart, E. 1998. Effects of mutations in the *Saccharomyces cerevisiae* RNA14 gene on the abundance and polyadenylation of its transcripts. *Mol. Gen. Genet.* 258:16-25.
- 15 McMahon, S.B., H.A. Van Buskirk, K.A. Dugan, T.D. Copeland, and M.D. Cole. 1998. The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell*. 94:363-74.
- 20 Murray, J.A.H., M. Scarpa, N. Rossi, G. Cesareni. 1987. Antagonistic controls regulate copy number of the yeast 2 μ plasmid. *EMBO J.* 6:4205-4212.
- 25 Nasr, F., A. Bécam, P.P. Slonimski, and C.J. Herbert. 1994. YBR1012, an essential gene from *S. cerevisiae*: construction of an RNA-antisense conditional allele and isolation of a multicopy suppressor. *CR Acad. Sci. Paris*. 317:607-613
- 30 Nasr, F., A. Bécam, S.C. Brown, D. De Nay. P.P. Slonimski, and C.J. Herbert. 1995. Artificial antisense regulation of YBR1012 an essential gene from *S. cerevisiae* which is important for progression through G1/S. *Mol. Gen. Genet.* 249:51-57.
- 35

- 34 -

- Nomura, T., N. Fujita, A. Ishihama. 1985. Promoter selectivity of *E. coli* RNA polymerase: analysis of the promoter system of convergently-transcribed *dnaQ-rnh* genes. *Nucleic Acids Research*. 13(21):7647-61.
- 5 Orlando, C., P. Pinzani, and M. Pazzagli. 1998. Developments in quantitative PCR. *Clin. Chem. Lab. Med.* 36(5):255-269.
- 10 Pla, J., C. Gil, F. Monteoliva, M. Sanchez, and C. Nombela. 1996. Understanding *Candida albicans* at the molecular level. *Yeast*. 12:1677-1702.
- 15 Reeder, R.H. and W.H. Lang. 1997. Terminating transcription in eukaryotes: lessons learned from RNA polymerase I. *Trends in Biochemical Sciences*. 22(12):473-7, 1997
- 20 Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Sinclair, D.A., L. Guarente. 1997. Extrachromosomal rDNA circles--a cause of aging in yeast. *Cell*. 25 91(7):1033-42.
- 30 Smith, V., D. Botstein, and P. O. Brown. 1995. Genetic footprinting: A genomic strategy for determining a gene's function given its sequence. *Proc. Natl. Acad. Sci. USA*. 92:6479-6483.
- 35 Stoesser, G., Moseley M.A., Sleep J., McGowran M., Garcia-Pastor M., Sterk P. 1998. *Nucleic Acids Res.* 26(1):8-15.
- Thompson-Jager, S. Domdey H. 1990. The intron of the yeast actin gene contains the promoter for an

- 35 -

antisense RNA. Current Genetics. 17(3):269-73.

Thompson, J.R., E. Register, J. Curotto, M. Kurtz, and
R. Kelly. 1998. An improved protocol for the
5 preparation of yeast cells for transformation by
electroporation. Yeast. 14:565-571.

Thrash, C., A.T. Bankier, B.G. Barrell, and R.
Sternglanz. 1985. Proc. Natl. Acad. Sci. USA 82: 4374-
10 4378.

Van Duin, M., J. van Den Tol, J.H. Hoeijmakers, D.
Bootsma, I.P. Rupp, P. Reynolds, L. Prakash, and S.
Prakash. 1989. Conserved pattern of antisense
15 overlapping transcription in the homologous human
ERCC-1 and yeast RAD10 DNA repair gene regions.
Molecular & Cellular Biology. 9(4):1794-8.

Wach, A., A. Brachat, R. Pohlmann, P. Philippsen.
20 1994. New heterologous modules for classical or PCR-
based gene disruptions in *Saccharomyces cerevisiae*.
Yeast. 10(13):1793-808.

Wilson, R.B., D. Davis, A.P. Mitchell. 1999. Rapid
25 hypothesis testing with *Candida albicans* through gene
disruption with short homology regions. Journal of
Bacteriology. 181(6):1868-74.

Zhu, J., W. Kempenaers, D. Van der Straeten, R.
30 Contreras, and W. Fiers. 1985. A method for fast and
pure DNA elution from agarose gels by centrifugal
filtration. Bio/Technology. 3: 1014-1016.

Table 1

Seq ID No.	Figure No.	Clone	Function
1	1	382c_cp	-
2	2	392c_cp	TUF1
3	3	-	RAD53
4	4	417c_cpG2L	-
5	5	323c_af	-
6	6	322c_cp ¹	-
7	7	26g3	-
8	8	409c_cp	-
9	9	382c_cpG1L2	-
10	14	382c_cp (prt)	-
11	15	392c_cp (prt)	TUF1
12	16		RAD53
13	17	325c_af (prt) ²	-
14	18	322c_cp (prt) ²	-
15	19	26g3 (prt)	-
16	20	417c_cp 92L (prt)	-

- 322c-cp is a member of the UPF0057 protein family. It contains potential transmembrane regions (6-23aa; 30-53aa) and could be low temperature or salt-stress inducible.
- 325c-af shows similarity to IMP4 yeast and related proteins and it might be involved in rRNA processing in *Candida albicans* in a similar way to IMP4.

- 37 -

Claims

1. A nucleic acid molecule encoding a polypeptide which is critical for survival and growth of the yeast *Candida albicans* and which nucleic acid molecule comprises any of the sequences of nucleotides illustrated in Seq ID Nos 1 to 9.

2. A nucleic acid molecule encoding a polypeptide which is critical for survival and growth of the yeast *Candida albicans* and which nucleic acid molecule comprises any of the sequences of nucleotides illustrated in Seq ID Nos 1 to 3.

3. A nucleic acid molecule encoding a polypeptide which is critical for survival and growth of the yeast *Candida albicans* and which nucleic acid molecule comprises any of the sequences of nucleotides illustrated in Seq ID Nos 1 or 2 and fragments or derivatives of said nucleic acid molecules.

4. A nucleic acid molecule according to any of claims 1 to 3 which is mRNA.

5. A nucleic acid molecule according to any of claims 1 to 3 which is DNA.

6. A nucleic acid molecule according to claim 5 which is cDNA.

7. A nucleic acid molecule capable of hybridising to the molecules according to any of claims 1 to 6 or the sequences illustrated in any of Seq ID Nos 1 to 9 under high stringency conditions.

8. An antisense molecule comprising a nucleic

- 38 -

acid molecule capable of hybridising to the molecules according to any of claims 1 to 6 or the sequences illustrated in any of Seq ID Nos 1 to 9.

5 9. Cells containing a nucleic acid molecule according to any of claims 1 to 8, wherein said cells are bacterial or eukaryotic.

10 10. A polypeptide encoded by the nucleic acid molecule according to any of claims 1 to 7 or the sequences illustrated in any of Seq ID Nos 1 to 9.

15 11. A polypeptide having any of amino acid sequences illustrated in any of Seq ID Nos 14 to 20.

12. A recombinant DNA construct comprising a nucleic acid molecule according to claim 5 or 6.

20 13. A recombinant DNA construct comprising a nucleic acid molecule according to claim 5 or 6 wherein said nucleic acid molecule is inserted in the antisense orientation.

25 14. A recombinant DNA construct according to claim 12 or 13 wherein said recombinant DNA construct is an expression vector.

30 15. A construct according to claim 14 which comprises an inducible promoter.

16. A construct according to claim 14 or 15 which comprises a sequence encoding a reporter molecule.

35 17. Cells containing a recombinant DNA construct according to any of claims 12 to 16, wherein said cells are bacterial or eukaryotic.

- 39 -

18. A nucleic acid molecule according to any of claims 1 to 8 or the nucleotide sequences illustrated in Seq ID Nos 1 to 9 for use as a medicament.

5 19. Use of a nucleic acid molecule according to any of claims 1 to 8 or the sequences illustrated in Seq ID Nos 1 to 9 in the preparation of a medicament for treating *Candida albicans* associated diseases.

10 20. A polypeptide according to claim 10 or 11 for use as a medicament.

 21. Use of a polypeptide according to claim 10 or 11 in the preparation of a medicament for treating
15 *Candida albicans* associated infections.

 22. A pharmaceutical composition comprising a nucleic acid molecule according to any of claims 1 to 8 or a polypeptide according to claim 10 or 11
20 together with a pharmaceutically acceptable carrier diluent or excipient therefor.

 23. A *Candida albicans* cell comprising an induced mutation in the DNA sequence encoding the
25 polypeptide according to claim 10.

 24. A method of identifying compounds which selectively modulate expression or functionality of polypeptides or metabolic pathways in which these
30 polypeptides are involved and which are crucial for growth and survival of *Candida albicans*, which method comprises:

 (a) contacting a compound to be tested with one or more *Candida albicans* cells having a
35 mutation in a nucleic acid molecule according to any of claims 1 to 8 which

- 40 -

mutation results in overexpression or underexpression of said polypeptides in addition to contacting one or more wild type *Candida albicans* cells with said compound,

5 (b) monitoring the growth and/or activity of said mutated cell compared to said wild type; wherein differential growth or activity of said one or more mutated *Candida* cells is indicative of selective action of

10 said compound on a polypeptide or another polypeptide in the same or a parallel pathway.

25. A compound identifiable according to the

15 method of claim 24.

26. A compound according to claim 25 for use as a medicament.

20 27. Use of a compound according to claim 25 in the preparation of a medicament for treating *Candida albicans* associated diseases.

28. A pharmaceutical composition comprising a

25 compound according to claim 25 together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

29. A method of identifying DNA sequences from a

30 cell or organism which DNA encodes polypeptides which are critical for growth or survival of said cell or organism, which method comprises:

(a) preparing a cDNA or genomic library from

35 said cell or organism in a suitable expression vector which vector is such that it can either integrate into the genome in said cell or that it permits transcription

- 41 -

of antisense RNA from the nucleotide sequences in said cDNA or genomic library.

- (b) selecting transformants exhibiting impaired growth and determining the nucleotide sequence of the cDNA or genomic sequence from the library included in the vector from said transformant.

30. A method according to claim 29 wherein said cell or organism is a yeast or filamentous fungus.

31. A method according to claim 29 or 30 wherein said cell or organism is any of *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* or *Candida albicans*.

32. An antibody capable of binding to a polypeptide according to claim 10 or 11.

33. An oligonucleotide comprising a fragment of from 10 to 120 contiguous nucleotides of a nucleic acid molecule according to any of claims 1 to 8.

34. An oligonucleotide according to claim 33 comprising a fragment of from 10 to 50 contiguous nucleotides.

*1/53**FIG. 1.*

AACAGCTGGT CTTCTGCTAA TACATTCAAC CCTTTCCATA TCTATACTCC
1 50

AACAATATGA TAACTGATGA ACAATTGAAT ACCATTGCAT TGACATTTGG
51 100

TTTTGCTTCA ATAATATTAA TCATAATATA TCATGCCATA TCTACTAATG
101 150

TACATAAATT AGAAGATGAA ACCCCATCAT CTTCATTTAC CAGAACAAAT
151 200

ACTACTGAAA CTACTGTTGC AAGTAAGAAA AAGAAGTAAT AACTGATGGA
201 250

TTTTTCTTCC TACCACCAAT TGAATAATGC TAGACTTGTT GGTGTGCTAC
251 300

AAATATTTCA AAAGAAAATA CGAATACTTT ATAAAATGGT AAGAACGGAA
301 350

GATGGTTTCT CATTTATACA CTAAATACAA ATCACATACA CATAACAAA
351 400

CACAAATACA TACATACACC TATATCCCTT TATTTGAT
401 438

2/53

FIG. 2.

ATGTTAAAAA CACTAACTCA AACTTTACGC TTAAGTGGGA AAGCTTTCCC
1 50

AAAGGTCCGT CCGGCCTTGA TCAGAACCTA CGCTGCCTTC GACCGTTCTA
51 100

AACCTCATGT CAACATTGGT ACTATTGGTC ATGTTGATCA TGGTAAACT
101 150

ACATTGACTG CTGCTATCAC CAAAGTTTTA GCCGAACAAG GTGGTGCCAA
151 200

CTTCTTGGAT TATGGTTCTA TTGATAGAGC TCCAGAAGAA AGAGCTAGAG
201 250

GTATCACTAT TTCCACTGCC CACGTTGAAT ACGAAACCAA GAACAGACAC
251 300

TATGCCCACG TTGATTGTCC AGGACACGCT GATTATATCA AAAATATGAT
301 350

TACTGGTGCC GCTCAAATGG ATGGTGCTAT CATTGTTGTT GCTGCCACTG
351 400

ATGGTCAAAT GCCTCAAACC AGAGAACATT TGTTATTGGC CAGACAAGTT
401 450

GGTGTTC AAG ACTTGGTTGT GTTTGTCAAC AAAGTCGATA CTATTGATGA
451 500

CCCTGAAATG TTGGAATTAG TCGAAATGGA AATGAGAGAA TTGTTATCCA
501 550

CCTACGGTTT TGATGGTGAC AACACTCCAG TTATTATGGG ATCTGCTTTA
551 600

ATGGCTTTGG AAGACAAGAA ACCAGAAATT GGTAAGGAAG CTATCTTGAA
601 650

ATTGTTAGAT GCTGTCGATG AACACATTCC AACTCCATCA AGAGACTTGG
651 700

AACAACCATT TTTGTTACCA GTTGAAGACG TGTTCTCCAT CTCCGGTAGA
701 750

GGAAGTGTG TCACTGGTAG AGTTGAAAGA GGTGTTTTGA AGAAGGGTGA
751 800

AGAAATCGAA ATTGTTGGTG GTTTTGACAA ACCTTACAAG ACTACTGTGA
801 850

CCGGTATTGA AATGTTCAAA AAAGAATTAG ACTCTGCTAT GGCTGGTGAC
851 900

*3/53**FIG. 2 (CONTINUED).*

AACTGTGGTG TTTTGTTAAG AGGTGTTAAA AGAGATGAAA TCAAGAGAGG
901 950

TATGGTTTTG GCCAAACCAG GTACTGCTAC TTCTCACAAG AAGTTCTTGG
951 1000

CTTCCTTGTA TATTTTGACT TCCGAAGAAG GTGGTCGTTC CACTCCATTT
1001 1050

GGTGAAGGTT ACAAGCCTCA ATGCTTCTTC AGAACTAACG ATGTCACTAC
1051 1100

CACATTTTCA TTCCCAGAAG GAGAAGGTGT TGATCATTCT CAAATGATCA
1101 1150

TGCCAGGTGA CAACATTGAA ATGGTTGGTG AATTGATCAA ATCTTGTCCA
1151 1200

TTAGAAGTCA ACCAACGTTT CAACTTGAGA GAAGGTGGTA AACTGTTGG
1201 1250

TACTGGTTTG ATTACCAGAA TCATCGAATA AACAGAATGT GCACTGTGAA
1251 1300

TAATAAAAAG AAAAGAGGTA TATATAGGTG ACTTTGTATT TTGTATTGAA
1301 1350

CAATAAAATT CTGTAAATAG TAAGGGCCTC
1351 1380

4/53

FIG. 3.

GAATTCGCCCTTAAGCACTCGTTTCAACTATACATTCAGTAACAACACCCTTAATTTACCAAACCTACA
TTAATGGAAGTA
ACACAACGGACGCAGAGTCAGACACAACCAACACAACAGTCACCGACAACCTCAGACGCAAACCCAAAG
CAAAGAGGACCA
GAATAGGATTTGTCAATTGATTTGCTCCACGGGTCAGTTTGGCAATTATGATTTGAATATCAACGATA
AAACTATCGTAC
AAGGTAAATGACGTGGTATTTTGAAGAGACCCCAACTCAGATTTGCAAGTGGCGTCGTCGTCGAGA
ATTTCAAACAAG
CATTTTCAAATCTGGCTCAACTTCAATGATAAATCACTATGGATAAAGGACACTTCAACTAACGGGAC
ACACCTTAACAA
CAGTCGATTGGTGAAAGGATCAAACCTACCTTCTTAATCAGGGTGATGAAATAGCAGTAGGGGTTGGTA
GAGACGAGGACG
TTGTGAGGTTTGTCTGTCTTTGGTGACAAATACAACCCGGCAAAGCTACCTGATTCGACCAACACA
ATTAAAGATGAA
GGAATATACAAAGACTTTATTGTGAAAAATGAAACGATAGGCCAAGGAGCATTGTCCTGTGAAAAA
GGCGATTGAACG
ATCTACGGGCGAGTCGTACGCGGTGAAGATTATAAATCGAAGAAAAGCATTAAATACCGGTGGTGAA
GTGCCATGGCAG
GAGTGGACCGTGAATTGTCCATATTAGAGCGGCTCAACCACCCAAATATAGTTGCTCTAAAAGCTTTT
TATGAAGATATG
GACAATTACTATATTGTGATGGAATTGGTGCCGGGCGGTGATTTGATGGACTTTGTGGCTGCAAACGG
TGCAATAGGAGA
AGACGCAACACAAGTGATCACGAAACAGATTCTAGAAGGAATTGCCTATGTTTATAATTTAGGAATCT
CCCATCGTGATT
TGAAGCCAGATAATATTTTGATTATGCAAGATGACCCAATACTTGTTAAAATCACCGACTTTGGATTG
GCAAAATTCAGT
GACAATCTGACGTTTATGAAAACCTTTTGTGGTACATTGGCGTATGTTGCTCCCGAAGTTATCACCGG
TAAGTATGGATC
ATCGCAGATGGAACGCAACAAAAGGACAACCTACTCTTCCTTGGTTGACATTTGGTCTTTGGGATGTT
TGGTTTATGTAC
TTTTAACTTCTCATTACCATTCAACGGGAAAACCAGCAACAAATGTTTGCCAAGATCAAAAGGGGC
GAATTCATGAG
GCTCCATTAAATTCATACGACATTTCTGAAGACGGAAGAGATTTCTTGCAAGTGTGCTGCTACAGGTTAA
TCCTAAACTAAG
GATGACGGCTGCTGAAGCTTTGAAACATAAATGGTTGCAAGACTTGTATGAAGAGGATTCTGTCAAAT
CATTGAGTTTAT
CGCAATCACAGTCGCAACAATCTCGAAAGATAGATAATGGTATCCATATCGAATCATTGAGCAAAATT
GATGAAGACGTT
ATGCTTCGTCCATTGGATAGCGAAAGAAATAGGAAATCAAGTAAACAGCAAGATTTCAAGGTACCCAA
GCGTGTGATTCC
GTTATCTCAACATCCTGCAACACCGTTACCAATGTCACAACCGAAAAAGAGGCCGTATCAAATAGACC
CTAGAACAAACA
AAAAAGTCGATTGGAAGAACCTCTGACAAGCAAGAAAGTCAAGCTAAGTGATTCCGTTGTTGCGGAA
GACTACTTGAAG
TTGGGGCCACTTGCAAATTCGTTATTCCAAGAAACAATAAATATTTCAAAGTCCCCGTTTCTTTCCG
AAGAAATGACAC
TTGTGATTGCGAGATAGACGACGACAGACTATCCAACTTCATTGTGTGATTACCAAAGAAAACGACT
CTATATGGTTAT
TGGATAAGAGTACTAACTCGTGCTTGGTCAACAATACTAGTGTTGGAAAAGGCAACAAAGTTTGTCTT
AGAGGAGGGGAG
ATATTACATCTCTTCTTTGACCCATTGTCACTGCAACATATAGGTTTCAAAGTAGTCCTTGTGATCA
ACTGTCTGGTGA
ACATAAGAGTCAAGTGGAGGTTTTGAAACAAACCTCAGAAGAAATGAATATTATTCCACTTATTTCTG
GTTTAAAGTAGTA
TAAGTTCATAGATTTAGCATATATACAAGCATTTCCTATAGAAACAAAGGTTTCAATTTAGTTATT
TACCTCCATGCA
ATTACATTTACTTCTTCTTCCAAGGGCGAATTCTGCAGATATC

5/53

FIG. 4.

ATGGGTAGTATGTGAAGATACAATATTGAAAGTGTCTACTAGAATATCTAAGATGTTTGAGCCCATGG
AC
ATTTTTGGATTTGATAATTAAAAAAGTAGCAATAGATTATTGCGTTGGAGAAAGAATCACCATAGTT
GC
AAGATTTGATAGATGTTAAAATGTTACGCAGGCGAAAGATGTAACATCTCTTAAAGTAAGAAGAATA
TG
GACATGAATAAAAATAGATAGCACTATTTTGGAACCTGTTGAAGATATTAAAATAGAATGGGATTTCA
AC
ATAGATATTCAAAGTAACGAAACCTCACAATCAAATAAAAACAACAGTAATACTAACAATTCAATTTT
TA
TTTTTATAGAGGGTACTCCATCTTTAGGTAAACGTCACAACAAATCTCACACCTTATGTAACAGATGT
GG
CCGTCGTTTCATTCCACGTCCAAAAGAAGACCTGTTCTTCTTGTGGTTACCCAGCTGCTAAAATGAGAT
CT
CACAAC TGGGCCTTAAAAGCCAAAAGAAGAAGAACTACTGGTACCGGTAGAATGGCTTACTTGAAACA
CG
TTACCAGAAGATTCAAGAACGGTTTCCAAACTGGTGTTGCTAAAGCTCAAACCCCTTCCGCTTAAACT
AA
TTACTGAAGTTATTGGTCATGCATTAGTCATTATTCAATTAAAGTCATGTTAAGCATAGCAAAGGAAGA
AT
TGGTTAGATTCTTGTTTAAAATGTAATGACTATTTAATATCTGTTTAAATAAGAGGTTTAGTCTTTAT
TT
TTTACGTATACACCAAAAAAAAAAAGAAACAAATAAAATCTGTATATTAATGTTGG

6/53

FIG. 5.

ATGGGTACTA GTACAAGTGA AGCATTGAAG AACATCAAAA ACAAACAGCG
1 50

AAGACAGAAA GTTTTTGCAG AAATAAAACA TGAAAAGAAT AAACAACGTC
51 100

ATAAGCAAAG AGCCGAAAGA GCTAAGGAAG AAAGAGAAAA CCCCGAATTA
101 150

AGAGAGGAAA GAATAGCAGC TAATATCCCA GATACTATAG ATAGCAAACG
151 200

TATTTATGAT GAGACTATAG CTGCTGAAGT TGAAGGAGAT GACGAGTTTC
201 250

AGTCATATTT CACTAACTTG TTGGAAGAAC CAAAGATTTT GTTGACAACA
251 300

AGTGCCAATG CTAAAAAACC GGCCTATGAA TTTGCAGACA TGATCATGGA
301 350

CTTTTTACCG AATGTGACAT TTATCAAAAG GAAGAAGGAA TATACAATGC
351 400

AAGATATGGC CAAATATTGC TCGAATAGAG ACTTCACTGC ATTGCTTGTC
401 450

ATCAACGAAG ACAAGAAGAA GGTCAATGGT ATAACGCTCA TCAATTTACC
451 500

TGAAGGGCCA ACATTTTATT TTTCGATTAC ATCAATAGTT GATGGGAAAA
501 550

GAATTAAGGG ACACGGGAAA GCTGGTGATT ATTTACCTGA GATTGTATTG
551 600

AATAATTTCA ATTCAAGATT GGGTAAAACT GTGGGAAGAC TATTTCAAAG
601 650

TATTTTCCCT CATAAACCTG AACTTCAAGG AAGACAAGTG ATTACTTTGC
651 700

ACAATCAACG TGATTATATT TTTTTCAGAA GACATAGATA TATTTTCAGA
701 750

AATGAGGAAA AGGTTGGATT GCAGGAATTG GGTCCGCAGT TTACATTAAA
751 800

GCTAAGAAGA ATGCAAAAGG GAGTACGTGG TGATGTTGTT TGGGAACACA
801 850

GACCAGATAT GGAAAGAGAT AAGAAGAAGT TTTATTTATA AGCGGGTGTA
851 900

*7/53**FIG. 5 (CONTINUED).*

TAAAGGTAGT AGTAGTGCGT TTATAAGTAT GTGTGTGTGT TTATGCATAG
901 950

ATGTGTAAAG AGTAATACAG CTAATTCG
951 978

8/53

FIG. 6.

AACTAATTTG TTAAACATC AATACCAAGA AGTTTTTACA ATTCAATCCC
1 50

ACATACACCA TTAATTATGA ATTCTGAAAA GATTATTGAA GTTATCATTG
51 100

CTATTTTCTT ACCACCAGTA GCTGTGTTTA TGAAATGTGG TGCCACTACC
101 150

CCATTATGGA TTAAGTTGGT ATTATGTATC TTTATTTGGT TCCCTGCTAT
151 200

CTTACATGCC TTATACGTTG TGTGAAAGA TTAAACAAAC ACCAGAGATT
201 250

TACTGCTTGA TGAATTGATT ACTCCAAAGA GTTGTGACTA GTTCCCAGTG
251 300

TGTTTTTTTT GCCTTCCAAC TTTCTTTTAC ATTTTTCAT TACTACCACT
301 350

GTCTTCCCCC CTATTTTGCA GAGTTTTCAA AATTTATCCA AAACATGTGA
351 400

GTCATTAAAC CATATTATTA TAATTATTCT TTTTGTATT TTTTCCCTT
401 450

AAAACACGTT AATTTATTAA TCGTTTCGTT GTTTGGTATT TTATTTTTTT
451 500

GTATTTATCA ATTGGAATAT ATATCTATAC ATGAATTTAT TATCCATTGT
501 550

ACCAATTGTT AAAACATTTT GTTAGTTTTT TGTTACTAGT ATAAAANNAT
551 600

AATAAAAGTT TANTTCAAC
601 619

9/53
FIG. 7.

ATGACATTAG GGTTCGATAA ATTCATAAGC AAGGTCAGCA CTCATAGACG
1 50

TCAATCTGAA CCATCAATCT TGGAAATCGC AGCCACCAAT TCTCAAAATA
51 100

AATCGAGAAG GCTAAGTATG GATAATGGTC ATTGTTATGT TCGTGAATCA
101 150

ACTAATAATC ATCATCATTT AAATACCGTC GTTGATAATT TACGACAGCG
151 200

TGCGGGATCG TTTTCATTTA TTTACATCA CCATAATCAC CATCAGAATA
201 250

GTCACGATAA TTATACTGTC GATCCCCTTA CATCAAACGG AGCACGAATT
251 300

TCCCGATCAC GTTCACGTT CAAATCAGTT GGGCACGGAG AAGCAATATC
301 350

ACCAGCGTAT TTTTCCAAGA ATAAAACCAA AGATTTAGTG AAACAGGAAA
351 400

CAGCACATAT CATTCTGAAG AAATTACTCA ACATGTTACA AGATTTGGAT
401 450

TTACAAAACC CTATTGCATT GAAAACAATA TCACAAGGTT CAGAATCAAA
451 500

GTTTTGTAAT ATCTACGTGT CTAACACTAA TAATTGTATT TACTTACCAG
501 550

CAGCAAGTTC AACAAGTTTC ACTTATGAAG ATGATGAAAA TGGCGGCGTT
551 600

ATAATTGCTG AAGATAGAAA TGATGAAATG CCAACAGCAG TTAATAACAA
601 650

TACTTTGTCA ATGGATAGTA TAAATCATTC AGAGACTGAT TTCCTGGATT
651 700

CTCCACCACC TCCAGATTTA TTTTCTAAAA TGAAATCATT CCATTACCA
701 750

AATTACTTGA CTTCAAAAAT CGATTCTGAA TGTCCAATTC CACATACATT
751 800

TGCTGTGATT GTTGAATTAA CCAAGGACTC TTTGATTATT AAAGATCTTC
801 850

ATTTCOAATT TCAGTCATTA ACTACCATTT TATGGCCAAC TGGGGATGCA
851 900

10/53

FIG. 7 (CONTINUED 1).

TATAATCGGA CTCATGCCAA GGAGAAATTT ACCATTGGGA ATATGGAATG
901 950

GCGTACATCT TTAAGCGACG CCGACTATTA TATCAATAGT TCTAATTCCA
951 1000

ACGATGTTAA GCTGAAAAAC TTGGGTCCTG AAGATCTTAT TAATCGAACT
1001 1050

AGAGAATACA AATTAATCGA TATTGAAGAA CCAAACAATT CATCAAACAG
1051 1100

TTTACTGGAT GATGACATGG ATATTAATAA TATTACGTCG CCATTATCAA
1101 1150

CGTCACCAAC ATCAAGTTCA ACTTCAACAA ATTCAACCTC CAACTCATTG
1151 1200

GGTTCAGATT CATATAAAGC TGGTCTTTAT GTATTTTTAT TACCAATCTT
1201 1250

ATTGCCAGAA CATATTCCTG CTTCCATTGT TTCTATTAAT GGTTCATTGG
1251 1300

CTCATACATT ACTGGTTGAA TGCAATAAAT AACTGATAA GTTGAATCGG
1301 1350

AAATCAAAAG TATCAGCATC GTACAATTTA CCTATGGTCC GTACTCCACC
1351 1400

AAACATTGGT AATTCCATTG CTGATAAGCC AATTTATGTT AATAGGATTT
1401 1450

GGAATGATGC CGTACATTAT ATTATAACTT TCCCCGCAA ATATGTTACT
1451 1500

TTGGGTTGTG AACACATGAT AAATGTGAAA TTAGTGCCCA TGGTGAAAGA
1501 1550

TGTGGTTATC AAGCGTATTA AATTTAATGT ATTGGAGAGA ATAACCTTATG
1551 1600

TTTCCAAAAA TTTATCACGA GAATATGATT ATGATAGTGA AGACCCCTAT
1601 1650

TGTATTCATC CAGTTTCTAA AGAAAATAAA GTACGTGAAC GTGTTGTGTC
1651 1700

GTTATATGAA TTGAAAACGA AGGCAAAACA ATCTTCTGGT GGACATCTTG
1701 1750

AAGCTTATAA ACAAGAAGTT ATGAAATGTC CGGAAAATAA CCTTTTATTT
1751 1800

11/53

FIG. 7 (CONTINUED 2).

TCTTGTTATG AGGTTGAAAA TGATAATAAT AACGGCAACG GCAACGGCAA
1801 1850

CGGCAACGGA AACAGAACG TTAAACAAAA GAATAAAGAT CAACCAATGA
1851 1900

TTGCTACACC TTTAGATATC AATGTTTCTT TACCATTTTT AACTACTATG
1901 1950

TCTGATAGTT TAATTATGAC ATCAGCCATA GAAGAAGAAG GTTCAGATCT
1951 2000

GCCTCATACA TCAAGAAGAG GGTCGGCAGT GAGTATGACT GATAATAATA
2001 2050

CTACCCCAAG TAACAATAAC CCTTTATCTC CATTTTTGGG AGCAGTGGAA
2051 2100

ACTAATGGTG CTAGTATAAA TGAAATTGGT GATCATACAT TATTCCTGA
2101 2150

TTCTAATTTT CGACATATTG AAATTAAACA TCGATTACAA GTTACATTTA
2151 2200

GGATTTCTAA ACCGGATCTG GATAATAAAA TGCATCATTA TGAAGTGGTT
2201 2250

ATTGATACCC CCATCGTTTT ACTTAGTTCA AAATGTCAAG AAGATTCTCC
2251 2300

TCCTCCTTAT AGTTCTGTA
2301 2319

*12/53**FIG. 8.*

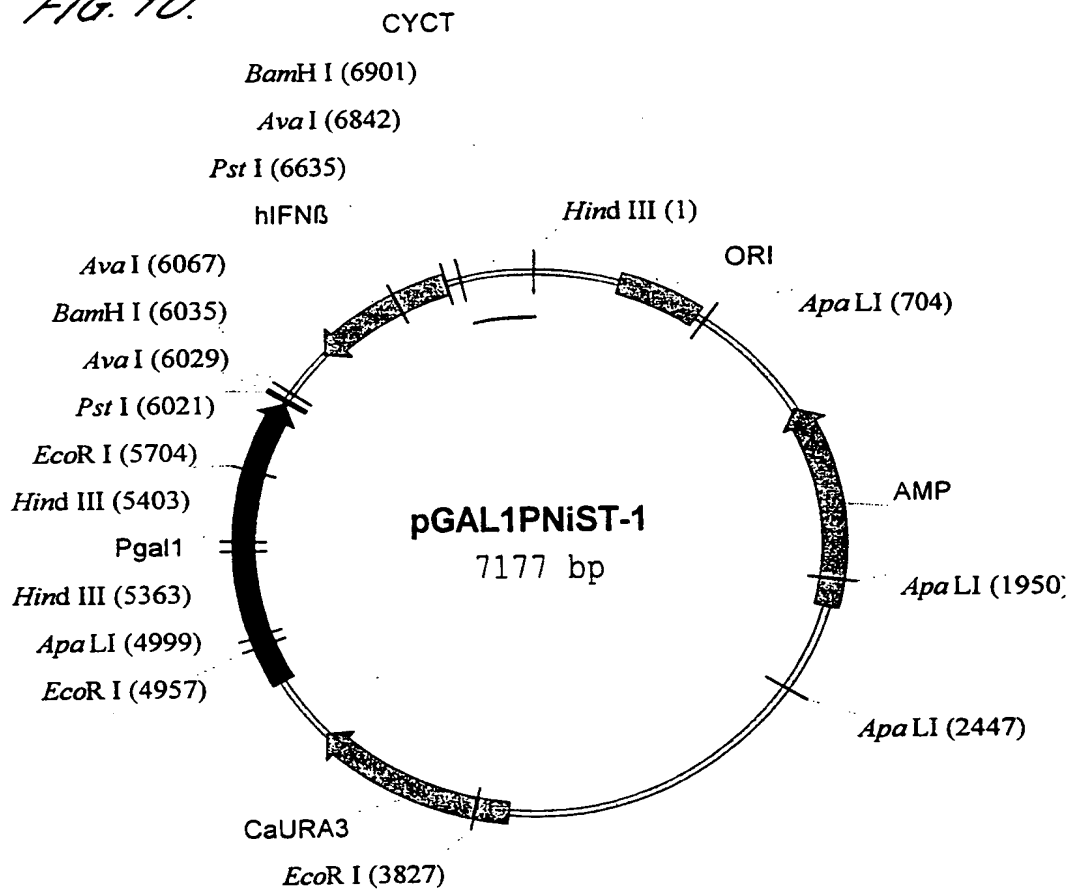
AACGTTTCGTG CAAAAGGCTA TACTGGTGAT ATCCACGCAG ATGAAGAGCA
1 50
AGTTTAATCA ACTCTTTGTC AATTAATGCT GTACTTGTTT TCATTTTATT
51 100
TGCTGGCATT TAAAGAATAC CCATAGTTCA GAAAATAAAA TTGAAAAATT
101 150
TAAAAAATAA CGCAATATCA TTCATTTTTT TTGTTTTTTT GACAATAATA
151 200
TTAATATGTA GTTACCAATG TTTTATGATT TTATATGTTT TGAAAAAATA
201 250
GTTTG
251

FIG. 9.

AACCTTACAA TCATTATACC AACTATCAAA ATCATAAGAC TCTTNACTT
1 50
CTGTTTTTGA TAGTTGGTAT AATGATTTAT GTATTATCTT AATTCATTAT
51 100
TATTAGTTTC GGTCACAAA
101 119

13/53

FIG. 10.



14/53
FIG. 11.

1 AGCTTGAGTA TTCTATAGTG TCACCTAAAT AGCTTGGCGT AATCATGGTC
51 ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAATT CCACACAACA
101 TACGAGCCGG AAGCATAAAG TGTAAAGCCT GGGGTGCCTA ATGAGTGAGC
151 TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTTCC AGTCGGGAAA
201 CCTGTGCTGC CAGCTGCATT AATGAATCGG CCAACGCGCG GGGAGAGGCG
251 GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC
301 TCGGTGCTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT
351 ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA
401 AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT
451 TTCCATAGGC TCCGCCCCC TGACGAGCAT CACAAAAATC GACGCTCAAG
501 TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC
551 CTGGAAGCTC CCTCGTGCGC TCTCCTGTTT CGACCCTGCC GCTTACCGGA
601 TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC
651 ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT
701 GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC
751 TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC CACTGGCAGC
801 AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
851 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAG GACAGTATTT
901 GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG
951 CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT TTTTTGTTT
1001 GCAAGCAGCA GATTACGCGC AGAAAAAAG GATCTCAAGA AGATCCTTTG
1051 ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT CACGTTAAGG
1101 GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA
1151 ATTAAAAATG AAGTTTTTAA TCAATCTAAA GTATATATGA GTAACTTGG
1201 TCTGACAGTT ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG
1251 TCTATTTCTG TCATCCATAG TTGCCTGACT CCCCGTCGTG TAGATAACTA
1301 CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA

15/53

FIG. 11 (CONTINUED 1).

1351 GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC AGCCAGCCGG
1401 AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
1451 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT
1501 TTGCGCAACG TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC
1551 GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA
1601 CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG
1651 ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA TGGTTATGGC
1701 AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
1751 TGA CTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA
1801 CCGAGTTGCT CTTGCCCCGGC GTCAATACGG GATAATACCG CGCCACATAG
1851 CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG GGGCGAAAAC
1901 TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT
1951 GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG TTTCTGGGTG
2001 AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
2051 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT
2101 TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA
2151 AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA GTGCCACCTG
2201 ACGTCTAAGA AACCATTATT ATCATGACAT TAACCTATAA AAATAGGCGT
2251 ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG GTGAAAACCT
2301 CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
2351 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT
2401 CGGGGCTGGC TTA ACTATGC GGCATCAGAG CAGATTGTAC TGAGAGTGCA
2451 CCATATGCGG TGTGAAATAC CGCACAGATG CGTAAGGAGA AAATACCGCA
2501 TCAGGCGAAA TTGTAAACGT TAATATTTTG TTAAAATTCT CGTTAAATAT
2551 TTGTTAAATC AGCTCATTTT TTAACCAATA GGCCGAAATC GGCAAAATCC
2601 CTTATAAATC AAAAGAATAG ACCGAGATAG GGTTGAGTGT TGTTCCAGTT
2651 TGGAACAAGA GTCCACTATT AAAGAACGTG GACTCCAACG TCAAAGGGCG

16/53

FIG. 11 (CONTINUED 2).

2701 AAAAACCGTC TATCAGGGCG ATGGCCCACT ACGTGAACCA TCACCCAAAT
2751 CAAGTTTTTTT GCGGTCGAGG TGCCGTAAAG CTCTAAATCG GAACCCTAAA
2801 GGGAGCCCCC GATTTAGAGC TTGACGGGGA AAGCCGGCGA ACGTGGCGAG
2851 AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG CGCTAGGGCG CTGGCAAGTG
2901 TAGCGGTCAC GCTGCGCGTA ACCACCACAC CCGCCGCGCT TAATGCGCCG
2951 CTACAGGGCG CGTCCATTCG CCATTCAGGC TCGCAACTG TTGGGAAGGG
3001 CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCTGGCGA AAGGGGGATG
3051 TGCTGCAAGG CGATTAAGTT GGGTAACGCC AGGGTTTTCC CAGTCACGAC
3101 GTTGTA AAAC GACGGCCAGT GAATTGTAAT ACGACTCACT ATAGGGCGAA
3151 TTGGTTTTCC AATGATGAGC ACTTTTAAAG TTCTGCTATG TGGCGCGGTA
3201 TTATCCCGTG TTGACGCCGG GCAAGAGCAA CTCGGTCGCC GCATACACTA
3251 TTCTCAGAAAT GACTTG GTT AGTACTAATA GGAATTGATT TGGATGGTAT
3301 AAACGGAAAC AAAAAAAGA GCTGGTACTA CTTTCTTTAA AATTATTTTA
3351 TTATTTGATT TTATTTAATA GTATATATTA TATTTTGAAC GTAGATTATT
3401 TTGTTGAAAG TTGCTGTAGT GCCATTGATT CGTAACACTA ATTCTGTATT
3451 AGTCATTCCT CTGTTTGAT AGTATCCAAA AAAACGGCTA TTTTTTTGCA
3501 ATCTTATTTT CTGCATATTA TACAGATAAC ATAATGAAAG AAAAAATCTT
3551 TTTTTTTGTT CTTCAATGAT GATTTCAACC ATTCTTTTAA ACATTGATCA
3601 ATTCCTGAGC AACAACCCCA TACACACTGG TTTATATACC GCCCCTTTAA
3651 CAGTTGAAGA AAGAAATAGA AATAGAAATA GCAAACAAAA GATATGACAG
3701 TCAACACTAA GACCTATAGT GAGAGAGCAG AAACATCATGC CTCACCACTA
3751 GCACAGCGAT TATTTTCGATT AATGGAACTG AAGAAAACCA ATTTATGTGC
3801 ATCAATTGAC GTTGATACCA CTAAGGAATT CCTTGAATTA ATTGATAAAT
3851 TAGGTCCTTA TGTATGCTTA ATCAAGACTC ATATTGATAT AATCAATGAT
3901 TTTTCCTATG AATCCACTAT TGAACCATTA TTAGAACTTT CACGTAAACA
3951 TCAATTTATG ATTTTTGAAG ATAGAAAATT TGCTGATATT GGTAATACCG
4001 TAAAGAAACA ATATATTGGT GGAGTTTATA AAATTAGTAG TTGGGCAGAT

17/53

FIG. 11 (CONTINUED 3).

4051 ATTACCAATG CTCATGGTGT CACTGGGAAT GGAGTGGTTG AAGGATTAAA
4101 ACAGGGAGCT AAAGAAACCA CCACCAACCA AGAGCCAAGA GGGTTATTGA
4151 TGTTAGCTGA ATTATCATCA GTGGGATCAT TAGCATATGG AGAATATTCT
4201 CAAAAAATG TTGAAATTGC TAAATCCGAT AAGGAATTTG TTATTGGATT
4251 TATTGCCCAA CGTGATATGG GTGGCCAAGA AGAAGGATTT GATTGGCTTA
4301 TTATGACACC TGGAGTTGGA TTAGATGATA AAGGTGATGG ATTAGGACAA
4351 CAATATAGAA CTGTTGATGA AGTTGTTAGC ACTGGAACTG ATATTATCAT
4401 TGTTGGTAGA GGATTGTTTG GTAAAGGAAG AGATCCAGAT ATTGAAGGTA
4451 AAAGGTATAG AAATGCTGGT TGGAATGCTT ATTTGAAAAA GACTGGCCAA
4501 TTATAAATGT GAAGGGGGAG ATTTTCACTT TATTAGATTT GTATATATGT
4551 AGAATAAATA AATAAATAAG TTAAATAAAT AATTAAATAA GGGTGGTAAT
4601 TATTACTATT TACAATCAAA GGTGGTCCTT CTAGCTGTAA TCCGGGCAGC
4651 GCAACGGAAC ATTCATCAGT GTAAAAATGG AATCAATAAA GCCCTGCGCA
4701 GCGCGCAGGG TCAGCCTGAA TACGCGTTTA ATGACCAGCA CAGTCGTGAT
4751 GGCAAGGTCA GAATAGCCCA AGTCGGCCGA GGGGCCTGTA CAGTGAGGGA
4801 AGATCTGATA TTGACGAAGA GGAACCAATG TAACGTTACA CTGAAGAAAA
4851 CACATAATAA ACGGGAAGAA ACGGTGTAAA AGTGTGAAAA TAATTTTTGA
4901 ATATCATTTT CCTTGGTTTA ATTCCAAACG AAACGTGTAT TTTTTTAGAG
4951 AATGGGAATT CTTATTGGAT GTCTAGATTG TTTGTTTACT CCAGACTGTG
5001 CACAAAAACG TTTGGATGGA TGATCAGAAG ATATTTTTAG GCTTAGCTCT
5051 AAATATAAGA AATGATGCTT GAAAATCCAG ACAGAAATTG AGTTTCAAAA
5101 ATTGGAATG TGAGGTATTA GTCAACTAAC CAAATAACAA TGCAAACCGG
5151 TTGATACATT TCATTTTGAA AATAATGAAA CTGGAATTGG ATGACCAGCA
5201 CACAAACACA TAAAGTAATT ATGGGAATTA GAAGCGAACA TAGAGGAATA
5251 CTTTGCCACG AACAGAATAC AAGTGGGAAC ACTTTTTTCT CCATTGTTTT
5301 AGTTCTGTTT TTTTGTCAAA CTGGTTTTGT GCTATGTGTA AAAAAATATT
5351 GCCAAGAAAA AAAGCTTGTT TTGTGGCCAG TGTCCGAAAA AAATTTTGGG

18/53

FIG. 14 (CONTINUED 4).

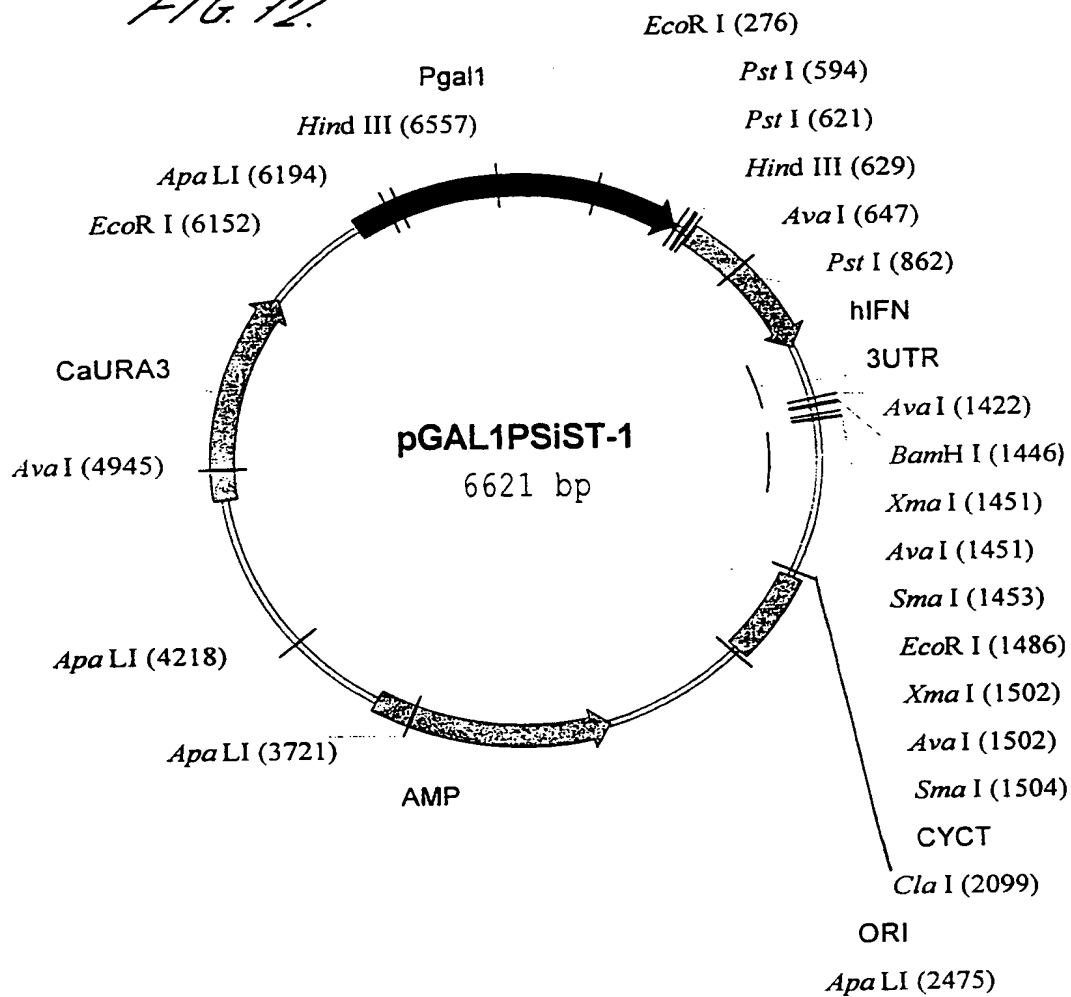
5401 GAAGCTTCGG ATTAATTTAT TTTTTTATTC CATCGGGGAA AGTGGGGGGG
5451 AAAAAAATT TAAGCAGTTC ATAAACCTT CAAAAAATA TATGGACAGA
5501 GATGATTGTA TTTTCCCGAC ACCAAAATCA TAATTAATA TGAGAAAATT
5551 GAATGTAACG TTACAATTTA TTTTATTTG AAGCTGAAAA GCGATTTATG
5601 ATTTTCCGA AATGAAAATT TTTTITAGGT TTATTTTTTT TGTCGGGGCAA
5651 AGAAAACTG AACAAGGATT ATTAAAATTT TTGGTGTGTTG TTTGTGTCTG
5701 GAGAATTCAT TCCTCTCTCA TCTTCACACA ATGTTTAGAC ATCTGACACG
5751 ATTCAAAATA GTTCGGTTTC CGGGGTGGT GTTAGTTTT CGTTTTTCGT
5801 TTTTTTGA AAGAATGTTT TAGCTCATTG GTTTCTTTC TTCATTCAAT
5851 AGTTTTGAAA GAATTTGCCC ACTTGTTATT ACAATCATAT AAAATTAAAC
5901 TTTGATATAA AATAGAGTTT GAAAGTTTCC CAGATCCTTT TTGATTTCCT
5951 TGTAATTTTT TTTCTCCCA CATATACACA CATAAAACC GATTTTTATA
6001 AGAAAGAGT ATACCCTGCA GCTCGACCTC GAGGGATCCG GGCCCTCTAG
6051 ATGCGGCCGC TAGGCCTCGA GGGACTTTTG CACCAAAAT AATTTATTTT
6101 CAAAATAAA ATTTAAATAA ATAAAATAA CTCATAATTT AATAAAAATT
6151 TCAAATCTT CTAGTGTCTT TTCATATGCA GTACATTAGC CATCAGTCAC
6201 TTAAACAGCA TCTGCTGGT GAAGAATGCT TGAAGCAATT GTCCAGTCCC
6251 AGAGGCACAG GCTAGGAGAT CTTAGTTTC GGAGGTAACC TGTAAGTCTG
6301 TTAATGAAGT AAAAGTTCCT TAGGATTTCC ACTCTGACTA TGGTCCAGGC
6351 ACAGTGACTG TACTCCTTGG CCTTCAGGTA ATGCAGAATC CTCCCATAAT
6401 ATCTTTTCAG GTGCAGACTG CTCATGAGTT TTCCCTGGT GAAATCTTCT
6451 TTCTCCAGT TTTCTTCCAG GACTGTCTTC AGATGGTTTA TCTGATGATA
6501 GACATTAGCC AGGAGGTTCT CAACAATAGT CTCATTCCAG CCAGTGCTAG
6551 ATGAATCTTG TCTGAAAATA GAAAGATGT TCTGGAGCAT CTCATAGATG
6601 GTCAATGCGG CGTCCTCCTT CTGGAAGTGC TGCAGCTGCT TAATCTCTC
6651 AGGGATGTCA AAGTTCATCC TGTCCTTGAG GCAGTATTCA AGCCTCCCAT
6701 TCAATTGCCA CAGGAGCTTC TGACACTGAA AATTGCTGCT TCTTTGTAGG

*19/53**FIG. 14 (CONTINUED 5).*

6751 AATCCAAGCA AGTTGTAGCT CATGGAAAGA GCTGTAGTGG AGAAGCACAA
6801 CAGGAGAGCA ATTTGGAGGA GACACTTGTT GGTCATGTTT CTCGAGGCCT
6851 TTTTGGCCAG CTGGCGCCTG CTGCGCGACG GCGAGCTGCT CACCACCCAG
6901 GATCCGTCCC CCTTTTCCTT TGTCGATATC ATGTAATTAG TTATGTCACG
6951 CTTACATTCA CGCCCTCCCC CCACATCCGC TCTAACCGAA AAGGAAGGAG
7001 TTAGACAACC TGAAGTCTAG GTCCCTATTT ATTTTTTTAT AGTTATGTTA
7051 GTATTAAGAA CGTTATTTAT ATTTCAAATT TTTCTTTTTT TTCTGTACAG
7101 ACGCGTGTAC GCATGTAACA TTATACTGAA AACCTTGCTT GAGAAGGTTT
7151 TGGGACGCTC GAAGGCTTTA ATTTGCA

20/53

FIG. 12.



21/53

FIG. 13.

1 TTCCATCGGG GAAAGTGGGG GGGAAAAAAT TTTAAGCAGT TCACAAAACC
51 TTCCAAAAAA TATATGGACA AAGATGATTG TATTTTCCCG ACACCAAAT
101 CATAATTAAT TATGAGAAAG TTAAATGTAA CGTTACAATT TATGTTTATT
151 TGAAGGTGAA AAGCGATTTA TGATTTTTC GAAATGAAAA TTTTTTTTAG
201 GTTTATTTTT TTTGTCGGGC AAAGAAAAAC TGAACAAGGA TTATTAAAAT
251 TTTTGGTGTT TGTTTGTGTC TGGAGAATTC ATTCCTCTCT CATCTTCACA
301 CAATGTTTAG ACATCTGACA CGATTCATGA TAGTTCGGTT TCCGGGGTTG
351 GTGTTTAGTT TTCGTTTTTC TTTTTTTTG GAAAGAATGT TTTAGCTCAT
401 TGGTTTTCTT TCTTCATTCA ATAGTTTGA AAGAATTTGC CCACTTGTTA
451 TTACAATCAT ATAAAATTAA ACTTTGATAT AAAATAGAGT TTGAAAGTTT
501 CCCAGATCCT TTTTGATTTC TTTGTAAATT TTTTTTTCTC CCACATATAC
551 ACACATACAA ACCGATTTTT ATAAGAAAGA GTTATACCCT GCAGCTCGAC
601 CTCGACTGTT TAAACCTGCA GGCATGCAAG CTTGGCCAAA AAGGCCTCGA
651 GGAACATGAC CAACAAGTGT CTCCTCCAAA TTGCTCTCCT GTTGTGCTTC
701 TCCACTACAG CTCTTCCAT GAGCTACAAC TTGCTTGAT TCCTACAAAG
751 AAGCAGCAAT TTTCAAGTGC AGAAGCTCCT GTGGCAATTG AATGGGAGGC
801 TTGAATACTG CCTCAAGGAC AGGATGAACT TTGACATCCC TGAGGAGATT
851 AAGCAGCTGC AGCAGTTCCA GAAGGAGGAC GCCGCATTGA CCATCTATGA
901 GATGCTCCAG AACATCTTTG CTATTTTCAG ACAAGATTCA TCTAGCACTG
951 GCTGGAATGA GACTATTGTT GAGAACCTCC TGGCTAATGT CTATCATCAG
1001 ATAAACCATC TGAAGACAGT CCTGGAAGAA AAAGTGGAGA AAGAAGATTT
1051 CACCAGGGGA AAAGTCATGA GCAGTCTGCA CCTGAAAAGA TATTATGGGA
1101 GGATTCTGCA TTACCTGAAG GCCAAGGAGT ACAGTCACTG TGCCTGGACC
1151 ATAGTCAGAG TGGAAATCCT AAGGAACTTT TACTTCATTA ACAGACTTAC
1201 AGGTTACCTC CGAAACTGAA GATCTCCTAG CCTGTGCCTC TGGGACTGGA
1251 CAATTGCTTC AAGCATCTT CAACCAGCAG ATGCTGTTTA AGTGACTGAT
1301 GGCTAATGTA CTGCATATGA AAGGACACTA GAAGATTTTG AAATTTTTAT

22/53

FIG. 13 (CONTINUED 1).

1351 TAAATTATGA GTTATTTTAA TTTATTTAAA TTTTATTTTG GAAAATAAAT
1401 TATTTTTTGGT GCAAAAGTCC CTCGAGGCCT AGCGGCCGCC TAGAGGATCC
1451 CCGGGCGCTA GCGGCCCGCT AGGCCTTTTT GGCCGAATTC GAGCTCGGTA
1501 CCCGGGGAGA TCCGTCCCCC TTTTCCTTTG TCGATATCAT GTAATTAGTT
1551 ATGTCACGCT TACATTCACG CCCTCCCCC ACATCCGCTC TAACCGAAAA
1601 GGAAGGAGTT AGACAACCTG AAGTCTAGGT CCCTATTTAT TTTTTTATAG
1651 TTATGTTAGT ATTAAGAACG TTATTTATAT TTCAAATTTT TCTTTTTTTT
1701 CTGTACAGAC GCGTGTACGC ATGTAACATT ATACTGAAAA CCTTGCTTGA
1751 GAAGGTTTTG GGACGCTCGA AGGCTTTAAT TTGCAAGCTA GCTTGCGCTA
1801 ATCATGGTCA TAGCTGTTTC CTGTGTGAAA TTGTTATCCG CTCACAATTC
1851 CACACAACAT ACGAGCCGGA AGCATAAAGT GTAAAGCCTG GGGTGCCTAA
1901 TGAGTGAGCT AACTCACATT AATGCGTTG CGCTCACTGC CCGCTTTCCA
1951 GTCGGGAAAC CTGTCGTGCC AGAGATCTCT GCATTAATGA ATCGGCCAAC
2001 GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTCCGC TTCCTCGCTC
2051 ACTGACTCGC TCGCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGATCG
2101 ATCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG
2151 CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA
2201 AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC CTGACGAGCA
2251 TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
2301 AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT
2351 CCGACCTGCG CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTGCGGAAG
2401 CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
2451 TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCGT TCAGCCCGAC
2501 CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA
2551 CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA
2601 GGTATGTAGG CCGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACTACGGC
2651 TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC

*23/53**FIG. 13 (CONTINUED 2)*

2701 CTTCCGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG
2751 GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA
2801 GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG ACGCTCAGTG
2851 GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA TCAAAAAGGA
2901 TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTAA ATCAATCTAA
2951 AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA
3001 GGCACCTATC TCAGCGATCT GTCTATTTCTG TTCATCCATA GTTGCCCTGAC
3051 TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC ATCTGGCCCC
3101 AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC CAGATTTATC
3151 AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT GGTCTTGCAA
3201 CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA AGCTAGAGTA
3251 AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTTGCCA TTGCTACAGG
3301 CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTC AGCTCCGGTT
3351 CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG CAAAAAAGCG
3401 GTTAGCTCCT TCGGTCCTCC GATCGTTGTC AGAAGTAAGT TGGCCGCACT
3451 GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTCATGC
3501 CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTC
3551 TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG CGTCAATACG
3601 GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA
3651 AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT GTTGAGATCC
3701 AGTTCGATGT AACCCTCTCG TGCACCCAAC TGATCTTCAG CATCTTTTAC
3751 TTTCAACCAGC GTTTCTGGGT GAGCAAAAAC AGGAAGGCAA AATGCCGCAA
3801 AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT ACTCTTCCTT
3851 TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA TGAGCGGATA
3901 CATATTTGAA TGTATTTAGA AAAATAAACA AATAGGGGTT CCGCGCACAT
3951 TTCCCCGAAA AGTGCCACCT GACGTCTAAG AAACCATAT TATCATGACA
4001 TTAACCTATA AAAATAGGCG TATCACGAGG CCCTTTCGTC TCGCGCGTTT

24/53

FIG. 13 (CONTINUED 3).

4051 CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA
4101 CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGGCGCG
4151 TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA
4201 GCAGATTGTA CTGAGAGTGC ACCATATCGA CGCTCTCCCT TATGCGACTC
4251 CTGCATTAGG AAGCAGCCCA GTAGTAGGTT GAGGCCGTTG AGCACCGCCG
4301 CCGCAAGGAA TGGTGCATGC AAGGAGATGG CGCCCAACAG TCCCCCGGCC
4351 ACGGGGCCTG CCACCATAACC CACGCCGAAA CAAGCACTAA TAGGAATTGA
4401 TTTGGATGGT ATAAACGGAA ACAAAAAAAAA GAGCTGGTAC TACTTTCTTT
4451 AAAATTATTT TATTATTTGA TTTTATTTAA TAGTATATAT TATATTTTGA
4501 ACGTAGATTA TTTTGTGAA AGTTGCTGTA GTGCCATTGA TTCGTAACAC
4551 TAATTCTGTA TTAGTCATTC CTCTTGTGTTG ATAGTATCCA AAAAAACGGC
4601 TATTTTTTTG CAATCTTATT TCCTGCATAT TATACAGATA ACATAATGAA
4651 AGAAAAAATC TTTTTTTTTG TTCTTCAATG ATGATTTCAA CCATTCTTTT
4701 AAACATTGAT CAATTCCTGA GCAACAACCC CATAACACT GGTTTATATA
4751 CCGCCCCTTT TACAGTTGAA GAAAGAAATA GAAATAGAAA TAGCAAACAA
4801 AAGATATGAC AGTCAACACT AAGACCTATA GTGAGAGAGC AGAAACTCAT
4851 GCCTCACCAG TAGCACAGCG ATTATTTCTGA TTAATGGAAC TGAAGAAAAC
4901 CAATTTATGT GCATCAATTG ACGTTGATAC CACTAAGGAG TTCCTCGAGT
4951 TAATTGATAA ATTAGGTCCT TATGTATGCT TAATCAAGAC TCATATTGAT
5001 ATAATCAATG ATTTTTCCTA TGAATCCACT ATTGAACCAT TATTAGAACT
5051 TTCACGTAAA CATCAATTTA TGATTTTTGA AGATAGAAAA TTTGCTGATA
5101 TTGGTAATAC CGTAAAGAAA CAATATATTG GTGGAGTTTA TAAAATTAGT
5151 AGTTGGGCAG ATATTACCAA TGCTCATGGT GTCAGTGGGA ATGGAGTGGT
5201 TGAAGGATTA AAACAGGGAG CTAAAGAAAC CACCACCAAC CAAGAGCCAA
5251 GAGGGTTATT GATGTTAGCT GAATTATCAT CAGTGGGATC ATTAGCATAT
5301 GGAGAATATT CTCAAAAAC TGTTGAAATT GCTAAATCCG ATAAGGAATT
5351 TGTTATTGGA TTTATTGCCC AACGTGATAT GGGTGGCCAA GAAGAAGGAT

*25/53**FIG. 13 (CONTINUED 4) .*

5401 TTGATTGGCT TATTATGACA CCTGGAGTTG GATTAGATGA TAAAGGTGAT
5451 GGATTAGGAC AACAAATATAG AACTGTTGAT GAAGTTGTGA GCACTGGAAC
5501 TGATATTATC ATTGTTGGTA GAGGATTGTT TGGTAAAGGA AGAGATCCAG
5551 ATATTGAAGG TAAAAGGTAT AGAAATGCTG GTTGAATGC TTATTTGAAA
5601 AAGACTGGCC AATTATAAAT GTGAAGGGGG AGATTTTCAC TTTATTAGAT
5651 TTGTATATAT GTAGAATAAA TAAATAAATA AGTTAAATAA ATAATTAAAT
5701 AAGGGTGGTA ATTATTACTA TTTACAATCA AAGGTGGTCC TTCTAGCTGT
5751 AATCCGGGCA GCGCAACGGA ACATTCATCA GTGTAAAAAT GGAATCAATA
5801 AAGCCCTGCG CTCATGAGCC CGAAGTGGCG AGCCCGATCT TCCCCATCGG
5851 TGATGTCGGC GATATAGGCG CCAGCAACCG CACCTGTGGC GCCGCAGCGC
5901 GCAGGGTCAG CCTGAATACG CGTTTAATGA CCAGCACAGT CGTGATGGCA
5951 AGGTCAGAAT AGCCCAAGTC GGCCGAGGGG CCTGTACAGT GAGGGAAGAT
6001 CTGATATTGA CGAAGAGGAA CCAATGTAAC GTTACACTGA AGAAAACACA
6051 CAATAAACGG GAAGAAACGG TGTAAGAGTG TGAAAATAAT TTTTGAATAT
6101 CATTTCCCTT GGTTTAATTC CAAACGAAAC GTGTTTTTTT TAGAGAATGG
6151 GAATTCTTAT TGGATGTCTA GATTGTTTGT TTA CTCCAGA CTGTGCACAA
6201 AAACGTTTGG ATGGATGATC AGAAGATATT TTTAGGCTTA GCTCTAAATA
6251 TAAGAAATGA TGCTTGAAAA ACCAGACAGA AATTGAGTTT CAAAAATTGG
6301 TAATGTGAGG TATTAGTCAA CTAACCAAAT AACAATGCAA ACCGGTTGAT
6351 ACATTTTCATT TTGAAAATAA TGAAACTGGA ATTGGATGAC CAGCACACAA
6401 ACACATAAAG TAATTATGGG AATTAGAAGC GAACATAGAG GAGTACTTGG
6451 CCACGAACAG AATACAAGTG GGAACACTAT TTTCTCCATT GTTTTAGTTC
6501 TGTTTTTTTG TCAGCCTAGT TTTGTGCTAT GTGTAAAAAA TATTGCCAAG
6551 AAAAAAAGCT TGTTTTGTGG CCAGTGTCCG AAAAAAATTT TGGGGAATCT
6601 TCGGATTAAT TTATGTTTTTC A

*26/53**FIG. 14.*

MITDEQLNTI ALTFGFASII LIIIIYHAIST NVHKLEDETP SSSFTRTNTT
1 50

ETTVASKKKK
51 60

FIG. 15.

MLKTLTQTLR LTGKAFFKVR PALIRTYAAF DRSKPHVNIG TIGHVDHGKT
1 50

TLTAAITKVL AEQGGANFLD YGSIDRAPEE RARGITISTA HVEYETKNRH
51 100

YAHVDCPGHA DYIKNMITGA AQMDGAIIVV AATDGQMPQT REHLLLARQV
101 150

GVQDLVVFN KVDTIDDPPEM LELVEMEMRE LLSTYGFDGD NTPVIMGSAL
151 200

MALEDKKPEI GKEAILKLLD AVDEHIPTPS RDLEQPFLLP VEDVFSISGR
201 250

GTVVTGRVER GVLKKGEEIE IVGGFDKPYK TTVTGIEMFK KELDSAMAGD
251 300

NCGVLLRGVK RDEIKRGMVL AKPGTATSHK KFLASLYILT SEEGGRSTPF
301 350

GEYKPCFF RTNDVTTTFS FPEGEGVDHS QMIMPGDNIE MVGELIKSCP
351 400

LEVNQRFNLR EGGKTVGTGL ITRIIE
401 426

27/53

FIG. 16.

MEVTQRTQSQ TOPTQQSPTT QTQTQSKEDQ NRICQLICST GQFGNYDLNI
1 50

NDKTIVQGKM TWYFGRDPNS DLQVASSSRI SNKHFQIWLN FNDKSLWIKD
51 100

TSTNGTHLNN SRLVKGSNYL LNQGDEIAVG VGRDEDVVRV VVVFQDKYNP
101 150

AKLPDSTNTI KDEGIYKDFI VKNETIGQGA FATVKKAIER STGESYAVKI
151 200

INRRKALNTG GGSAMAGVDR ELSILERLNH PNIVALKAFY EDMDNYYIYM
201 250

ELVPGGDLMD FVAANGAIGE DATQVITKQI LEGIAYVHNL GISHRDLKPD
251 300

NILIMQDDPI LVKITDFGLA KFSDNSTFMK TFCGTLAYVA PEVITGKYGS
301 350

SQMESQQKDN YSSLVDIWSL GCLVYVLLTS HLPFNGKNQQ QMFAKIKRGE
351 400

FHEAPLNSYD ISEDGRDFLQ CCLQVNPCLR MTAAEALKHK WLQDLYEEDS
401 450

VKSLSLSQSQ SQQSRKIDNG IHIESLSKID EDVMLRPLDS ERNRKSSKQQ
451 500

DFKVPKRVIP LSQHPATPLP MSQPKKRPYQ IDPRTNKKVD LEEPSTSKKV
501 550

KLSDSVVAED YLKLEPLANS LFQETINISK SPFSFGRNDT CDCEIDDDRL
551 600

SKLHCVITKE NDSIWLLDKS TNSCLVNNTS VGKGNKVLLR GGEILHLFFD
601 650

PLSSQHIGFK VVLVDQSSGE HKSQVEVLKQ TSEEMNIPL ISGLSSISS
651 699

*28/53**FIG. 17.*

MGTSTSEALK NIKNKQRRQK VFAEIKHEKN KQRHKQRAER AKEERENPEL
1 50

REERIAANIP DTIDSKRIYD ETIAAEVEGD DEFQSYFTNL LEEPKILLTT
51 100

SANAKKPAYE FADMIMDFLP NVTFIKRKKE YTMQDMAKYC SNRDFTALLV
101 150

INEDKKKVNG ITLINLPEGP TFYFSITSIV DGKRIKGHGK AGDYLPEIVL
151 200

NNFNSRLGKT VGRLFQSIFP HKPELQGRQV ITLHNQRDYI FFRRHRYIFR
201 250

NEEKVGLQE GPQFTLKLRRM QKGVRGDVVW EHRPDMERDK KKFYL
251 295

FIG. 18.

MNSEKIEVI IAIFLPPVAV FMKCGATTPL WINLVLCIFI WFPAILHALY
1 50

VVLKD
51

29/53
FIG. 19.

MTLGFDKFIS KVSTHRRQSE PSILEIAATN SQNKSRRLSM DNGHCYVRES
1
50
TNNHHHLNTV VDNLRQRAGS FSFISHHHNH HQNSHDNYTV DPLTSNGARI
51
100
SRSRSRSKSV GHGEAISPAY FSKNKTDLV KQETAHIISK KLLNMLQDLD
101
150
LQNPIALKTI SQGSESKFCK IYVSNTNNCI YLPAASSTSF TYEDDENGGV
151
200
IIAEDRNDDEM PTAVNNNTLS MDSINHSETD FSDSPPPPD L FSKMKSFHSP
201
250
NYLTSKIDSE CPIPHTFAVI VELTKDSLII KDLHFQFQSL TTILWPTGDA
251
300
YNRTHAKEKF TIGNMEWRTS LSDADYYINS SNSNDVSKN LGPEDLINRT
301
350
REYKLIDIEE PNNSSNSLSD DDMDINNITS PLSTSPTSSS TSTNSTSNL
351
400
GSDSYKAGLY VFLLPILLPE HIPASIVSIN GSLAHTLSVE CNKYTDKLN
401
450
KSKVSASYNL PMVRTPPNIG NSIADKPIYV NRIWNDVHY IITFPRKYVT
451
500
LGCEHMINVK LSPMVKDVI KRIKFNVLER ITYVSKNLSR EYDYDSEDPY
501
550
CIHPVSKENK VRERVVSLYE LKTKAQSSG GHLEAYKQEV MKCPENLLF
551
600
SCYEVENDDNN NGNGNGNGNG NKNVKQKNKD QPMIATPLDI NVSLPFLTTM
601
650
SDSLIMTSAI EEEGSDSPHT SRRGSAVSMT DNNTTPSNNN PLSPFLGAVE

*30/53**FIG. 19 (CONTINUED).*

651

700

TNGASINEIG DHTLFPDSNF RHIEIKHRLQ VTFRISKPDS DNKMHYEVV
701

750

IDTPIVLLSS KCQEDSPPPY SSV
751

773

*FIG. 20.*MGEOTPSLGKRHNKSHTLNRCGRRSFHVQKKTCS SCGYPAAKMRSHNWALKAKRRRTTGTGRMAYLK
HV
TRRFKNGFQTGVAKAQTPSA

31/53

FIG. 21.

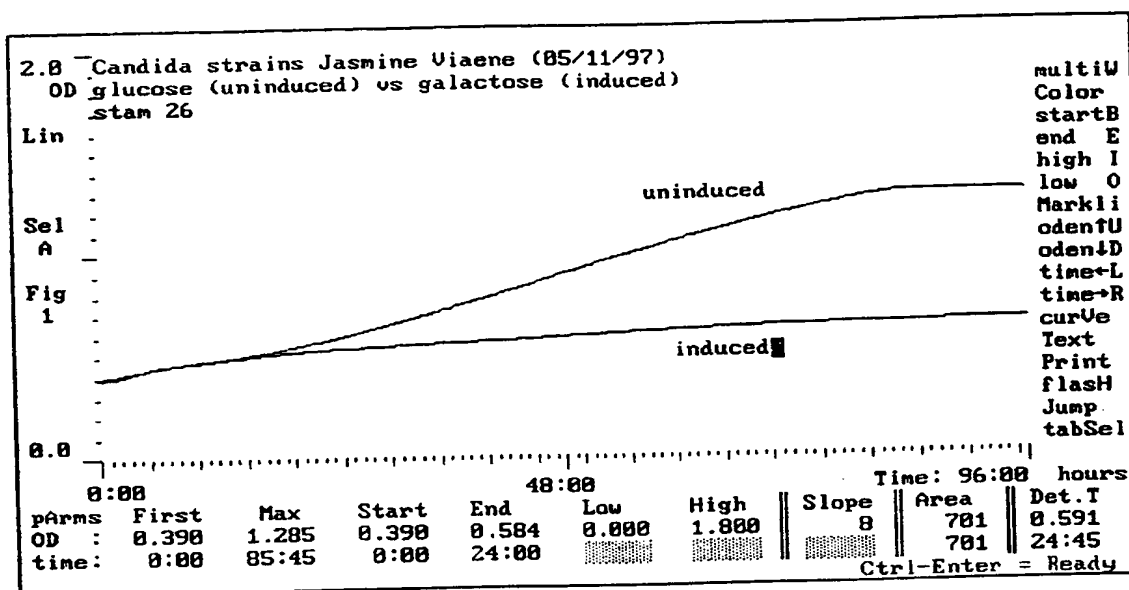
Project : Identification of novel essential genes in *C. albicans*

Strain no. : 26g

Freezer location : glycerol stocks box XXIII; C8

Growth curve(s) (Bioscreen) :

Date : 05/11/1997



Plasmid/clone name* : 26g3

Freezer location : original stocks box

Identifier (gene name) :

HTS screen :

Form generated by : Inge Loonen

(*) as it can be found in the *Candida albicans* Access dbase

32/53

FIG. 22.

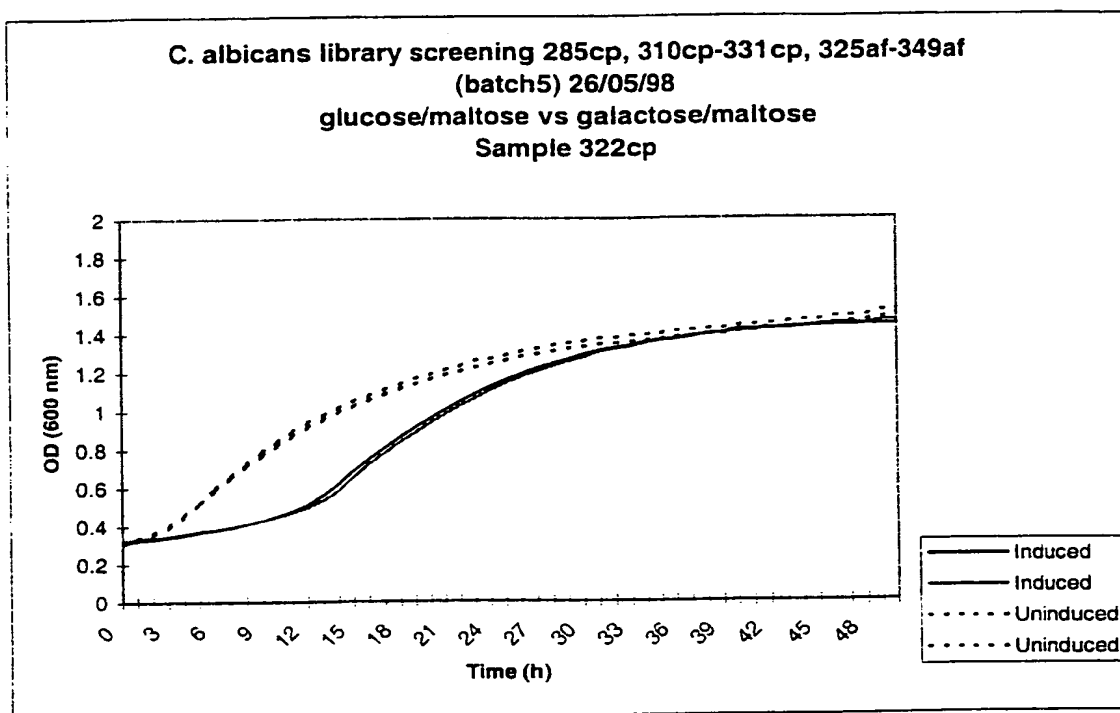
Project : Identification of novel essential genes in *C. albicans*

Strain no. : 322c_cp

Freezer location : glycerol stocks box XIV; D6

Growth curve(s) (Bioscreen) :

Date : 26/05/1998



Plasmid/clone name* : 322c_cp

Freezer location : original stocks box

Identifier (gene name) :

HTS screen :

Form generated by : Inge L onen

(*) as it can be found in the *Candida albicans* Access dbase

33/53

FIG. 23.

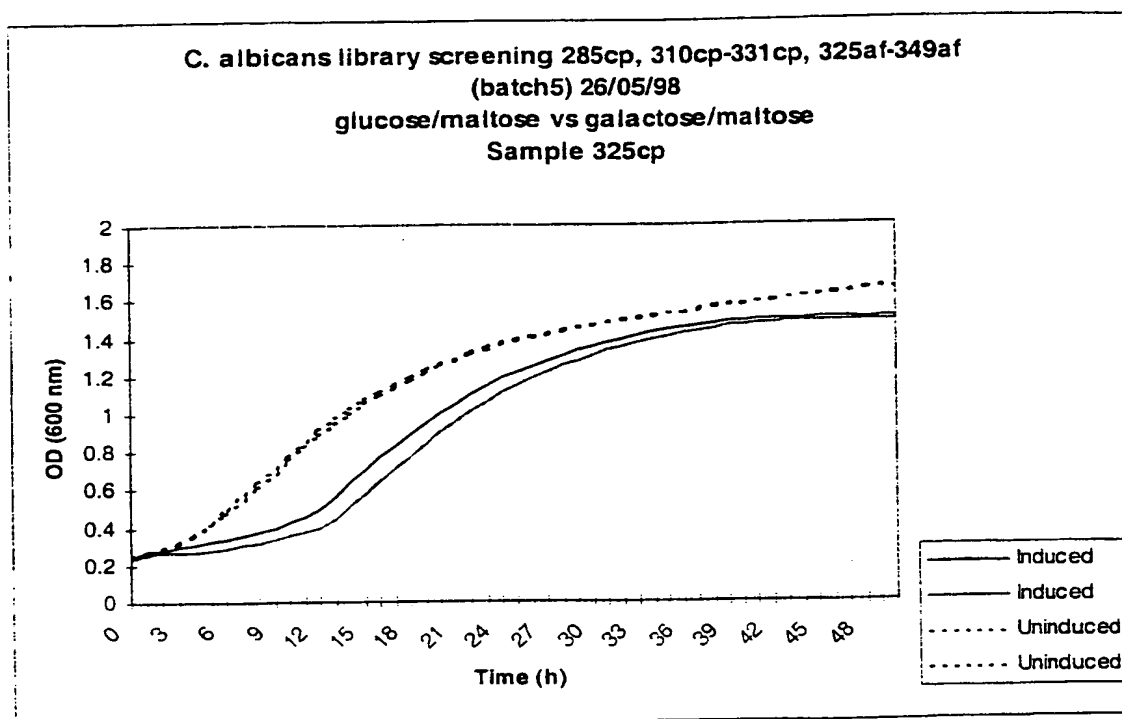
Project : Identification of novel essential genes in *C. albicans*

Strain no. : 325c_af

Freezer location : glycerol stocks box XIII; G4

Growth curve(s) (Bioscreen) :

Date : 26/05/1998



Plasmid/clone name* : 325c_af

Freezer location : original stocks box

Identifier (gene name) :

HTS screen :

Form generated by : Inge L onen

(*) as it can be found in the *Candida albicans* Access dbase

34/53

FIG. 24.

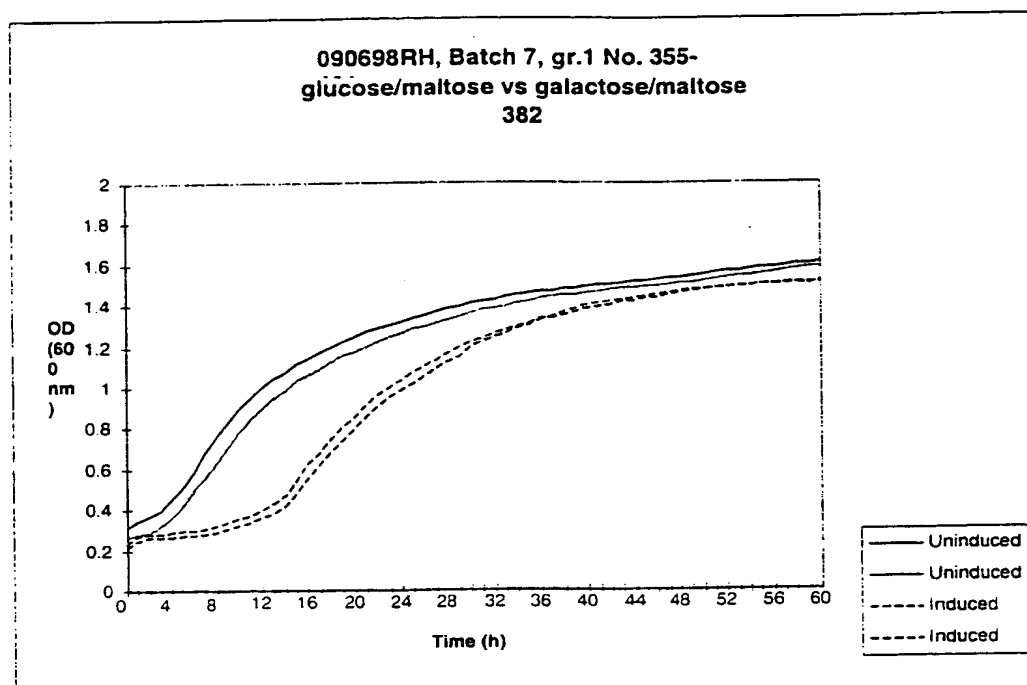
Project : Identification of novel essential genes in *C. albicans*

Strain no. : 382cp (FACS, batch 7, G1)

Freezer location : glycerol stocks box XVI; A2

Growth curve(s) (Bioscreen) :

Date : 09/06/98



Plasmid/clone name* : 382cp (purified PCR product)

Freezer location : original stocks box VIII; AAH8

Identifier (gene name) : OST4

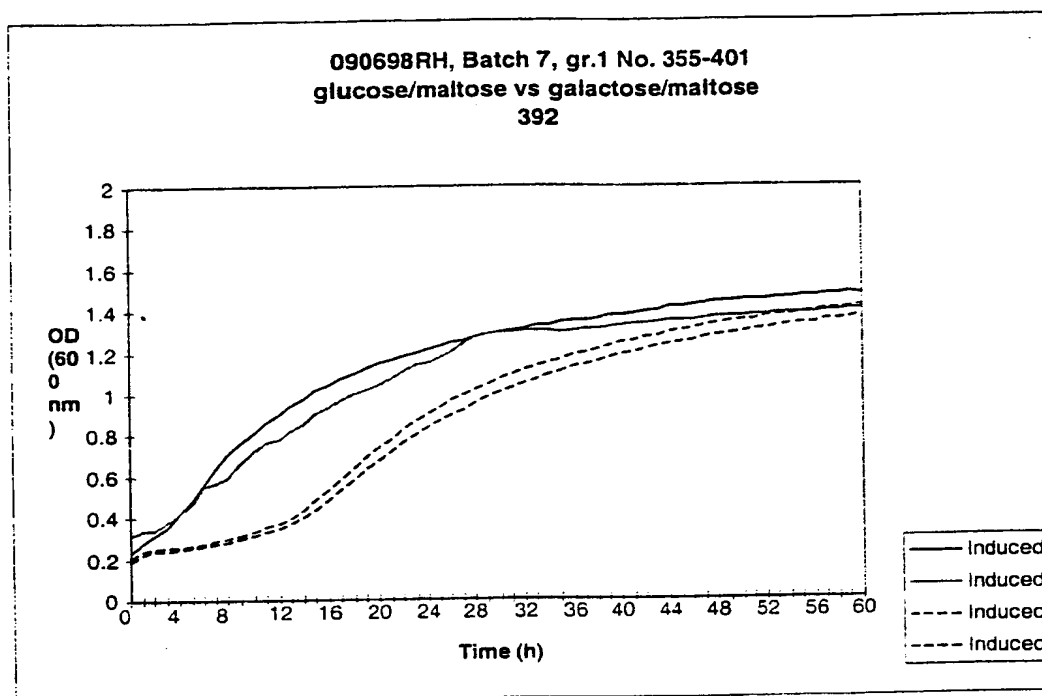
HTS screen :

Form generated by : Inge Loonen

(*) as it can be found in the *Candida albicans* Access dbase

35/53

FIG. 25.

Project : Identification of novel essential genes in *C. albicans***Strain no. : 392cp (FACS, batch 7, G1)****Freezer location : glycerol stocks box XVI; B3****Growth curve(s) (Bioscreen) :****Date : 09/06/98****Plasmid/clone name* : 392cp (purified PCR product)****Freezer location : original stocks box VIII; AAH2****Identifier (gene name) : TUF1****HTS screen :****Form generated by : Inge Loonen****(*) as it can be found in the *Candida albicans* Access dbase**

36/53

FIG. 26.

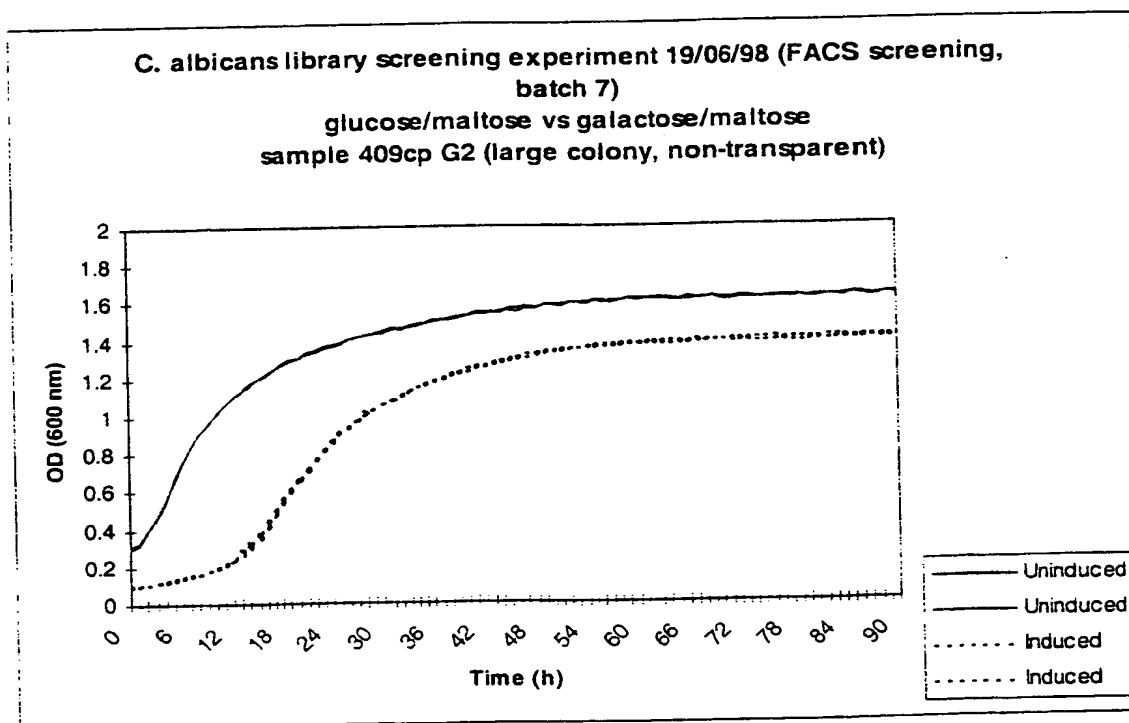
Project : Identification of novel essential genes in *C. albicans*

Strain no. : 409c_cp

Freezer location : glycerol stocks box XVI; C9

Growth curve(s) (Bioscreen) :

Date : 19/06/1998



Plasmid/clone name* : 409c_cp

Freezer location : original stocks box

Identifier (gene name) :

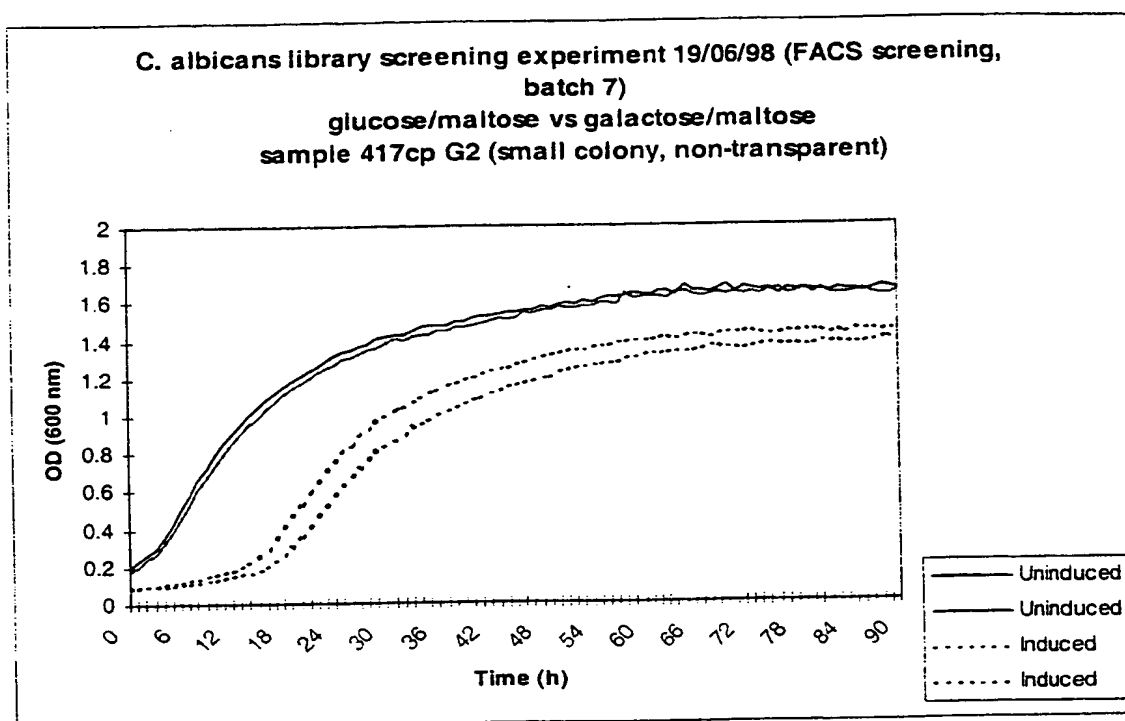
HTS screen :

Form generated by : Inge Loonen

(*) as it can be found in the *Candida albicans* Access dbase

37/53

FIG. 27.

Project : Identification of novel essential genes in *C. albicans***Strain no. : 417c_cpG2****Freezer location : glycerol stocks box XVI; D8****Growth curve(s) (Bioscreen) :****Date : 19/06/1998****Plasmid/clone name* : 417c_cpG2L****Freezer location : original stocks box****Identifier (gene name) :****HTS screen :****Form generated by : Inge Loonen****(*) as it can be found in the *Candida albicans* Access dbase**

38/53

FIG. 28.

Project : Identification of novel essential genes in *C. albicans*

Identifier (gene name) : 325caf

Disruptant strain :

Host strain :

Freezer location :

Disruption plasmid name* :

Freezer location :

Knock-out (single/double):

Lab book ref. :

Southern results :

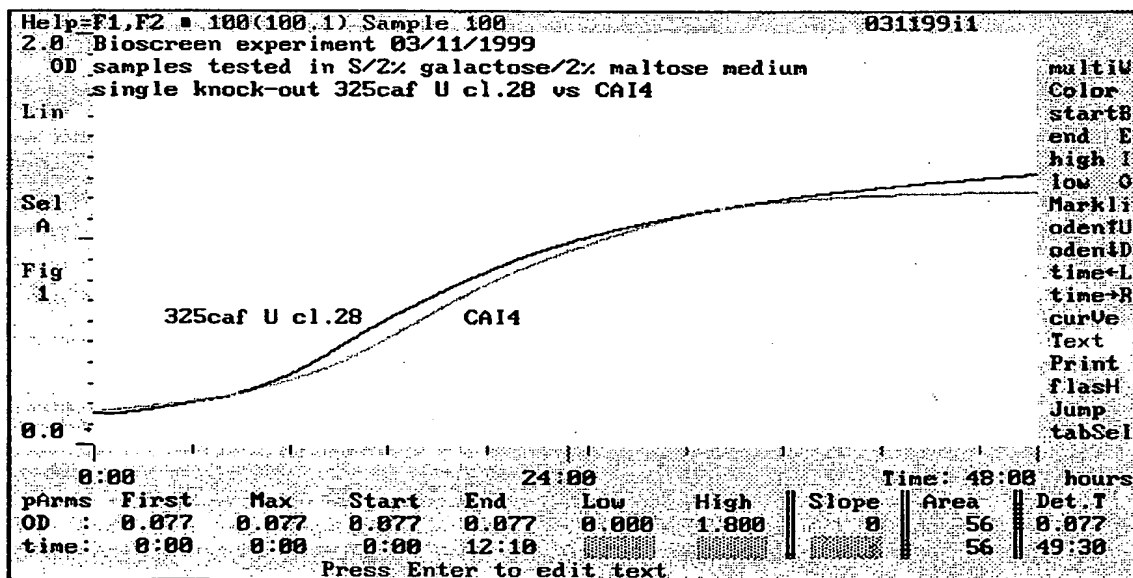
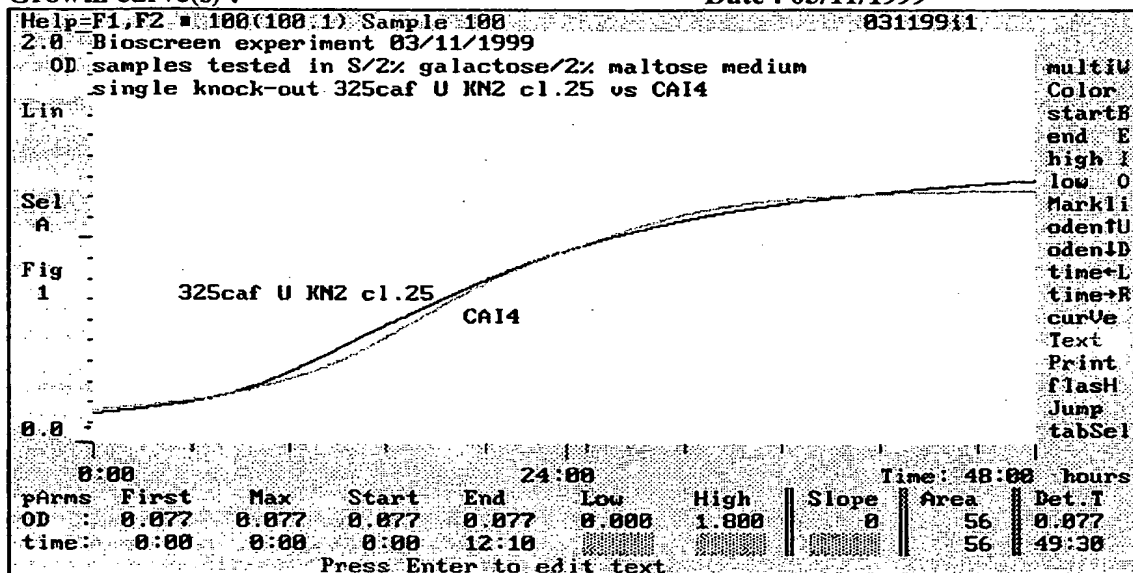
PCR results :

39/53

FIG. 28 (CONTINUED 1)

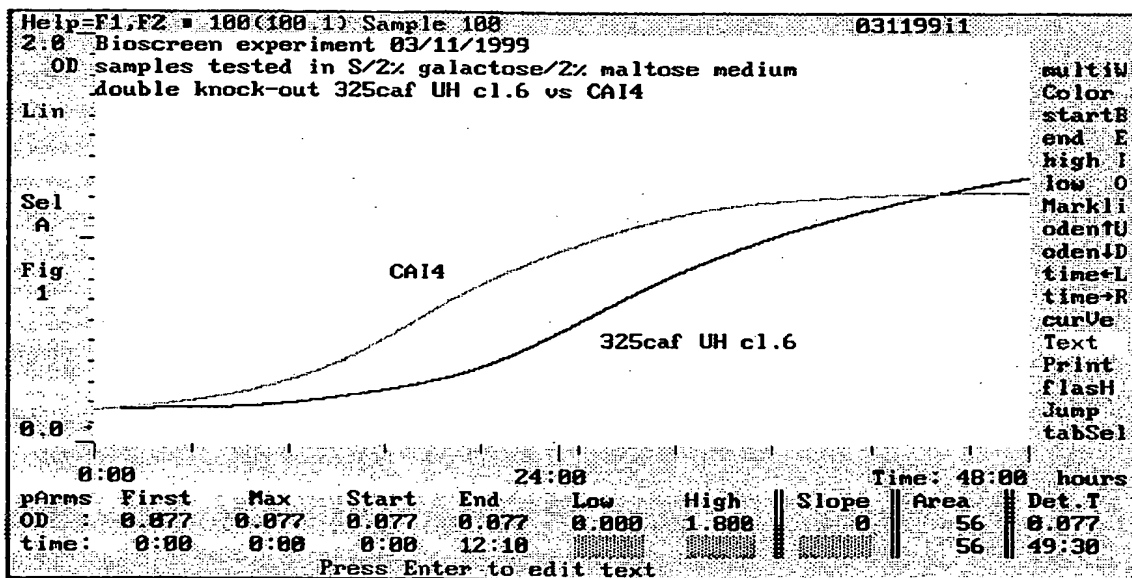
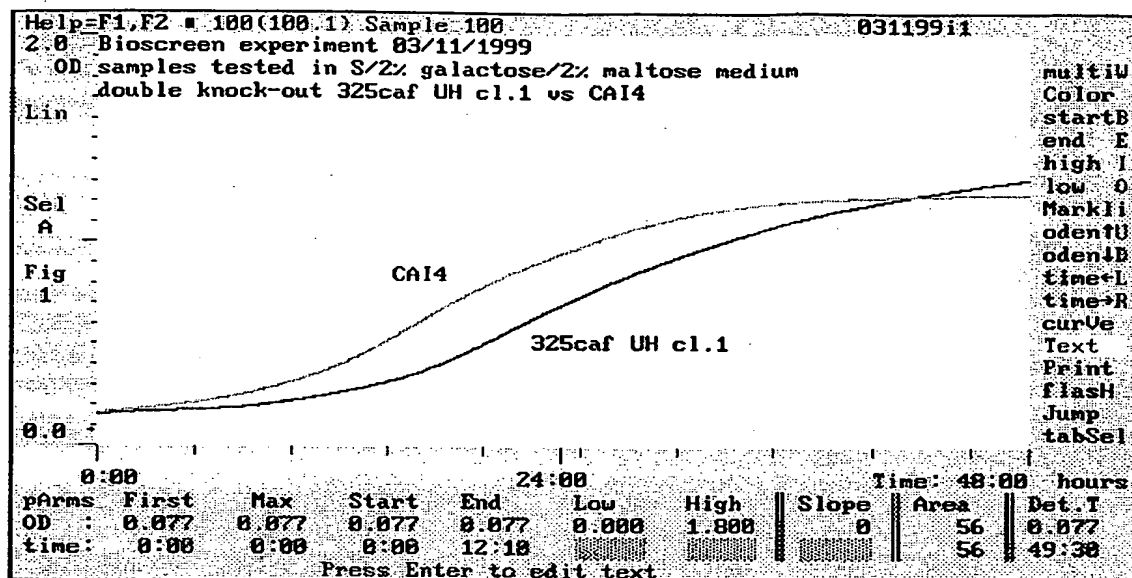
Growth curve(s) :

Date : 03/11/1999



40/53

FIG. 28 (CONTINUED 2)

**HTS screen :**

Bioscreen test of 325caf knock-out and WT growth in presence of hygromycin B

dilutions prepared

Stock solution of 53 mg/ml was prepared for hygromycin B.

From this solution dilutions were prepared of:

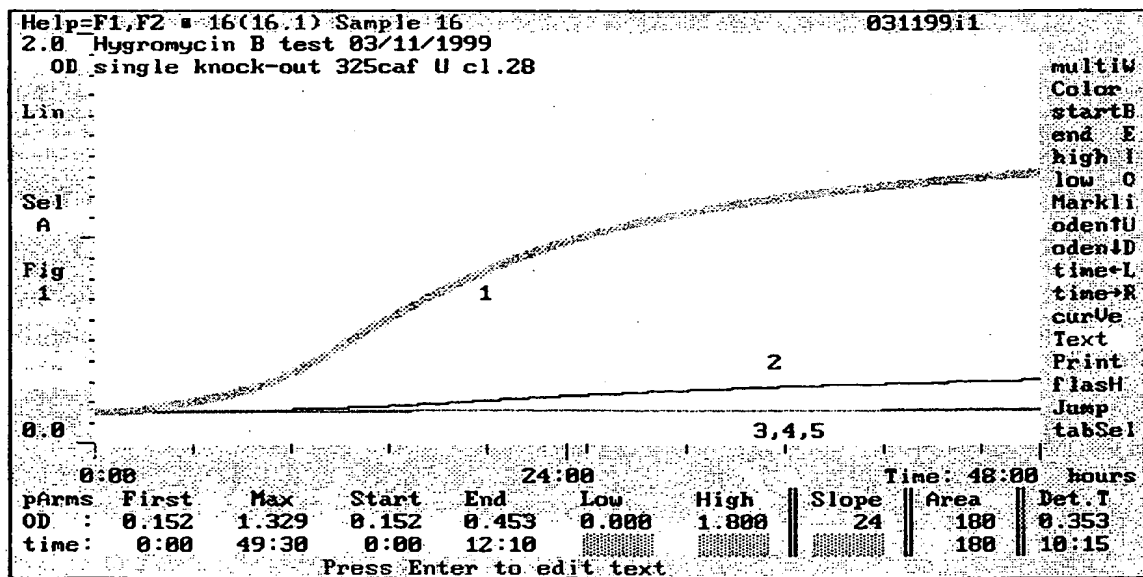
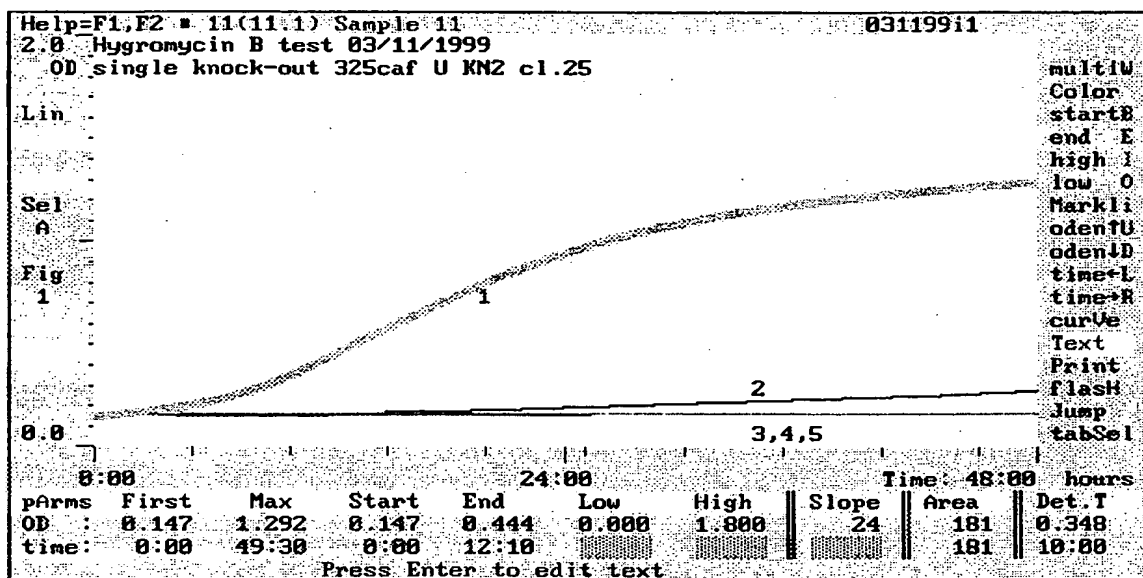
4000µg/ml, 3000µg/ml, 2000µg/ml and 1000µg/ml

Growth curves for 325cafG knock-out and WT in the presence of hygromycin B

SUBSTITUTE SHEET (RULE 26)

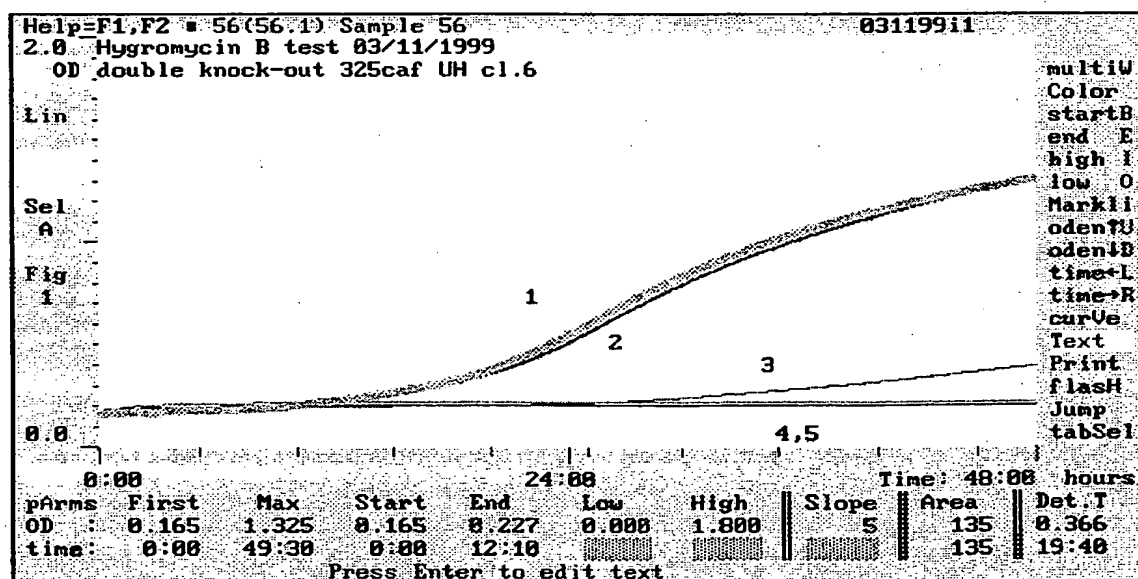
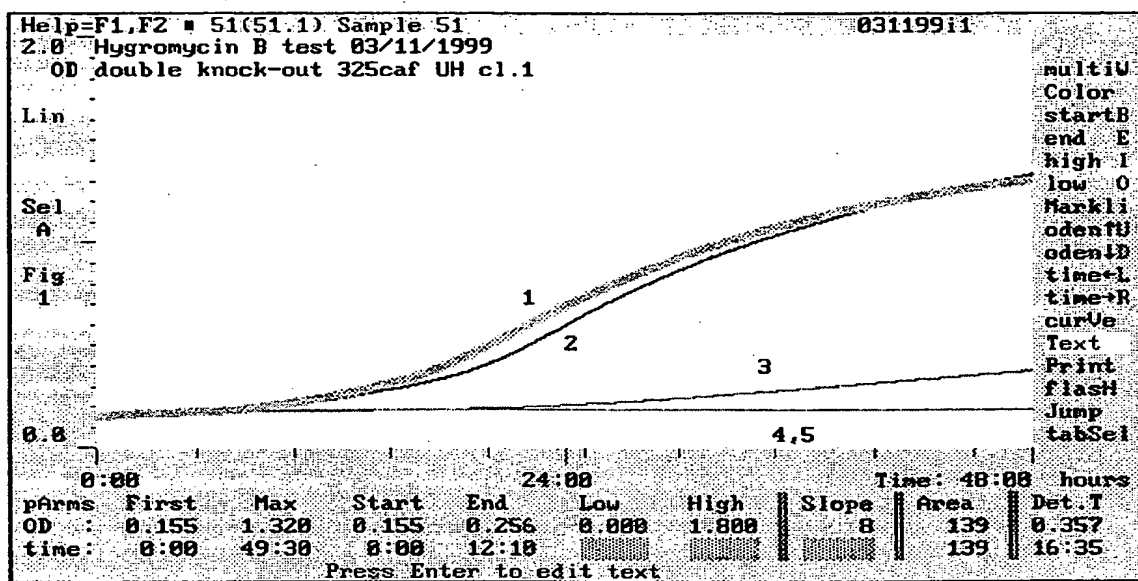
41/53

FIG. 28 (CONTINUED 3).



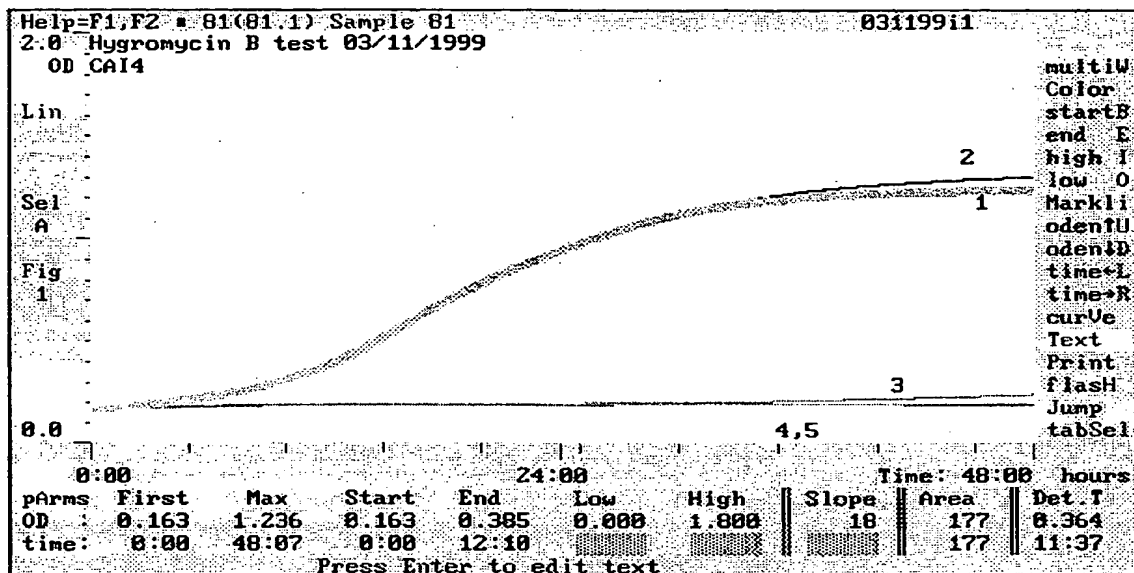
42/53

FIG. 28 (CONTINUED 4)



43/53

FIG. 28 (CONTINUED 5).

**Legend:**

- 1: S/2% gal/2% mal medium containing 0 µg/ml Hygromycin B
- 2: S/2% gal/2% mal medium containing 1000 µg/ml Hygromycin B
- 3: S/2% gal/2% mal medium containing 2000 µg/ml Hygromycin B
- 4: S/2% gal/2% mal medium containing 3000 µg/ml Hygromycin B
- 5: S/2% gal/2% mal medium containing 4000 µg/ml Hygromycin B

Form generated by :

(*) as it can be found in the Plasmid Access dbase

44/53

FIG. 29.

Project : Identification of novel essential genes in *C. albicans*

Identifier (gene name) : 322c_cp
Disruptant strain : 322c_cp (in progress)
Host strain : CAI4NG
Freezer location : Knockout strain, box, pos.

Disruption plasmid name* : 322c_cpURAcass.(inv)/pCR2.1(inv)
Freezer location : -

Knock-out (single/double): single (in progress)
Lab book ref. : Labbook 104 of Ronald de Hoogt

Southern results :

PCR results :

45/53

FIG. 30.

Project : Identification of novel essential genes in *C. albicans*

Identifier (gene name) : 417c_cpG2

Disruptant strain :

Host strain :

Freezer location :

Disruption plasmid name* :

Freezer location :

Knock-out (single/double):

Lab book ref. :

Southern results :

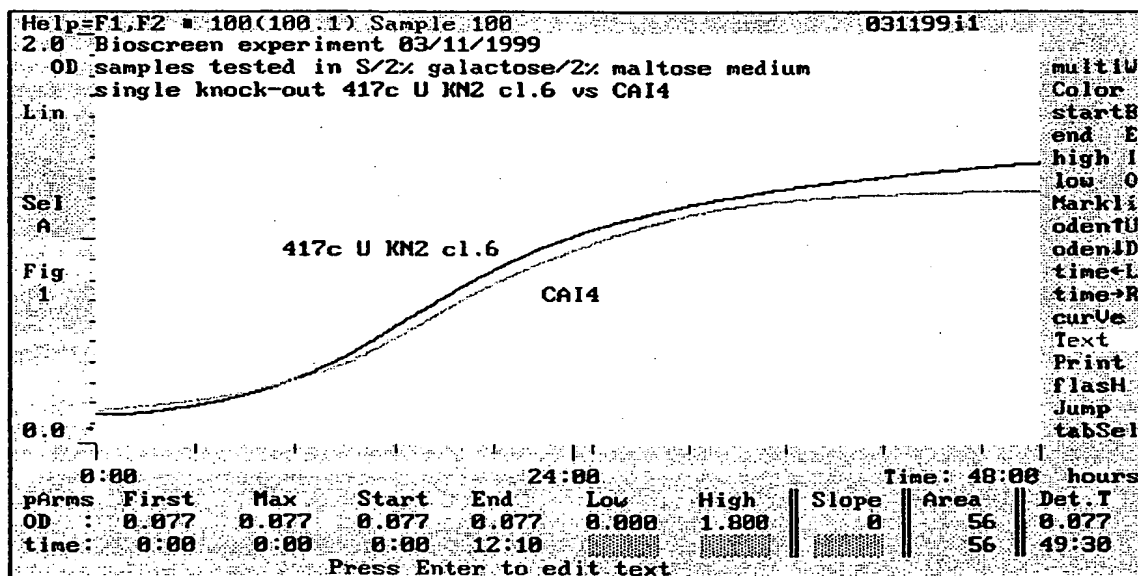
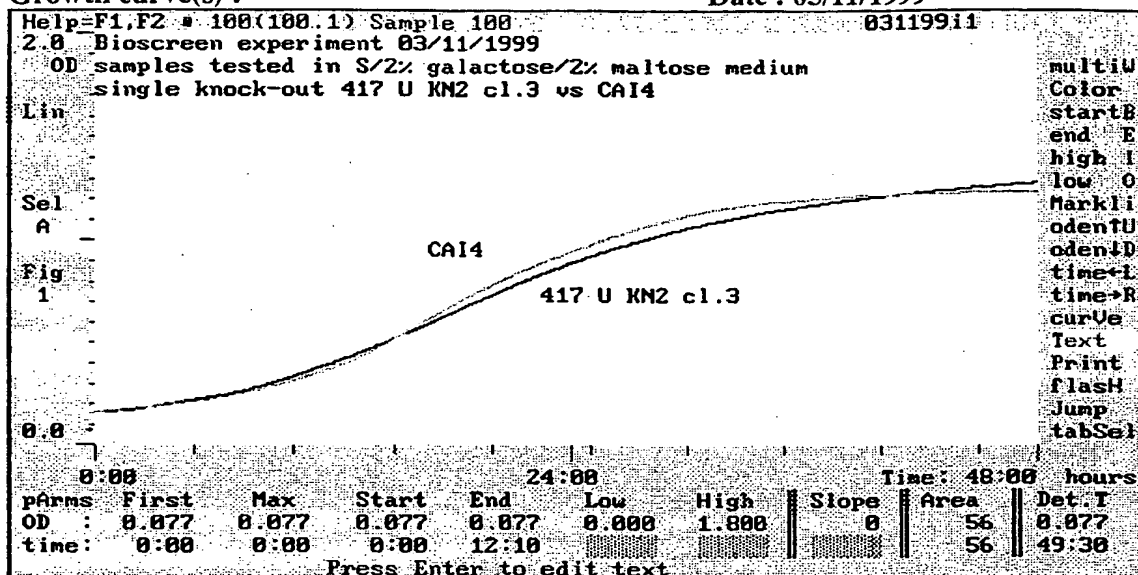
PCR results :

46/53

FIG. 30 (CONTINUED 1).

Growth curve(s) :

Date : 03/11/1999



HTS screen :

Bioscreen test of 417c_cp knock-out and WT growth in presence of hygromycin B

dilutions prepared

Stock solution of 53 mg/ml was prepared for hygromycin B.

From this solution dilutions were prepared of:

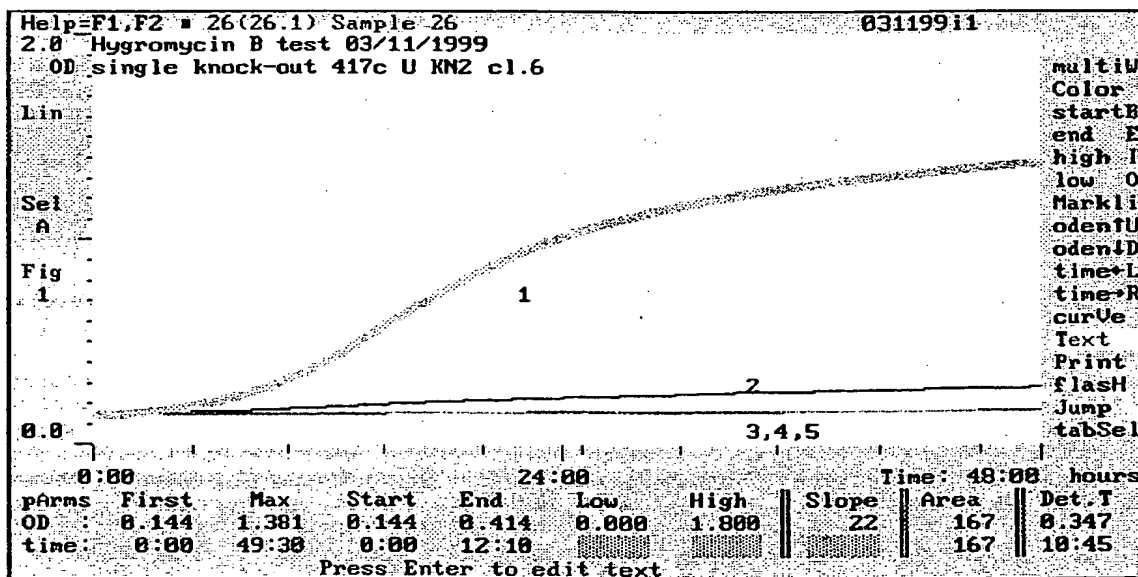
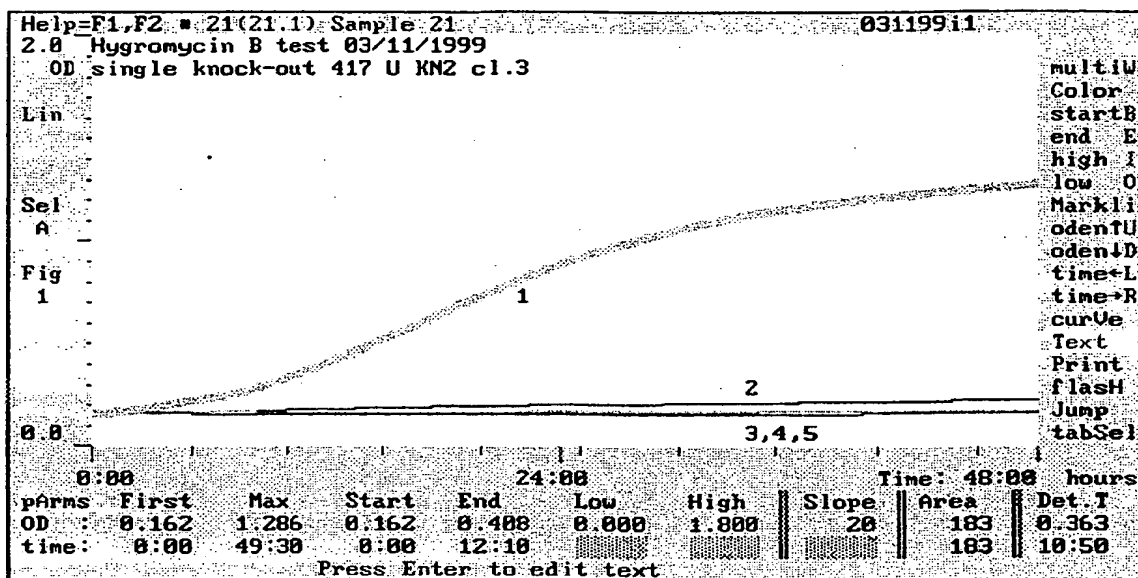
4000µg/ml, 3000µg/ml, 2000µg/ml and 1000µg/ml

Growth curves for 417c_cp knock-out and WT in the presence of hygromycin B

SUBSTITUTE SHEET (RULE 26)

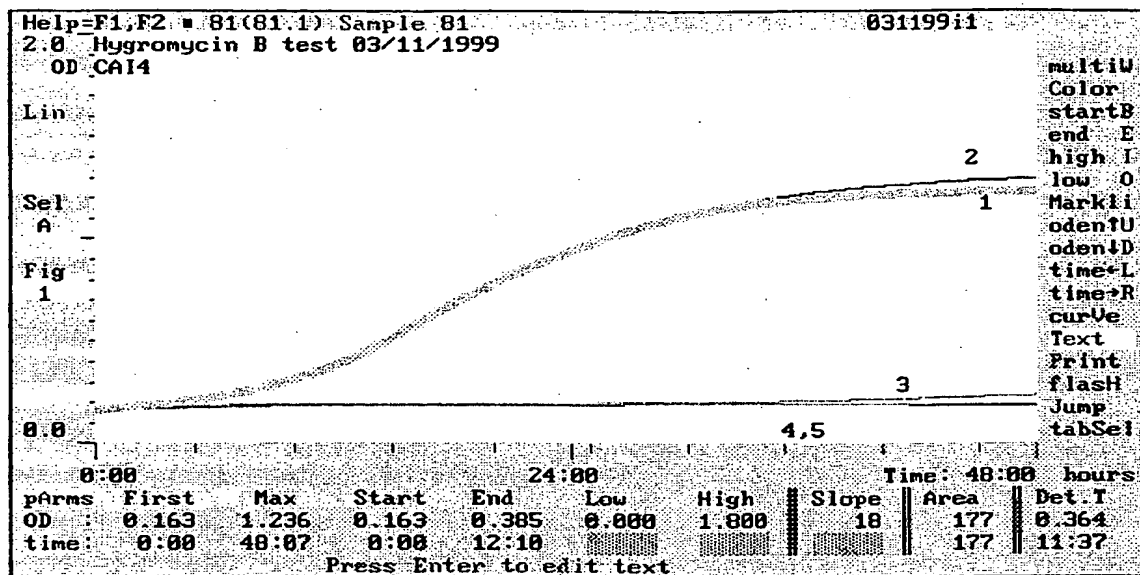
47/53

FIG. 30 (CONTINUED 2).



48/53

FIG. 30 (CONTINUED 3).

**Legend:**

- 1: S/2% gal/2% mal medium containing 0 μg/ml Hygromycin B
- 2: S/2% gal/2% mal medium containing 1000 μg/ml Hygromycin B
- 3: S/2% gal/2% mal medium containing 2000 μg/ml Hygromycin B
- 4: S/2% gal/2% mal medium containing 3000 μg/ml Hygromycin B
- 5: S/2% gal/2% mal medium containing 4000 μg/ml Hygromycin B

Form generated by :

49/53

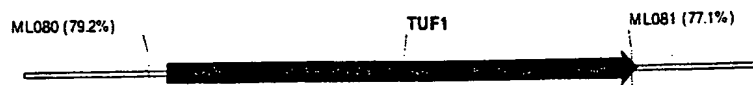
FIG. 31.

Project : Identification of novel essential genes in *C. albicans***Identifier (gene name) :** TUF1**Disruptant strain :** TUF1SAKO 7**Host strain :** CAI4/NG**Freezer location :** Strain collection Roland Contreras. YA132

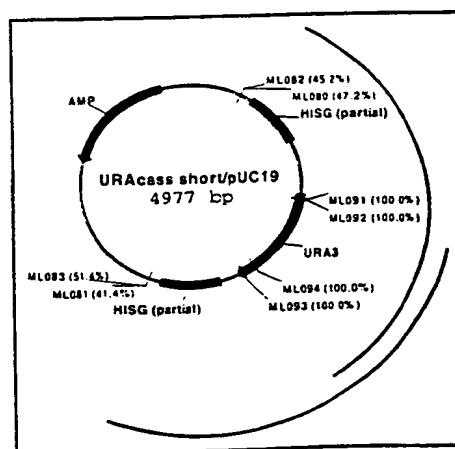
Disruption plasmid name* : Not applicable: short terminal homology (STH) PCR with overlapping fragments (split marker).

ML080	acataatcaagtgaatttacttacatcatttattgtggaaatttctgaatGTGCTGGAATTCGCCCTTTATG
ML081	tcacctatatataacctctttttctttttattattcacagtgcacattctgtCCGGCTCGTATGTGTGTG TGG
ML094	CCAGTGCTAACAACCTTCATCAACAGTT
ML092	GCCTCACCAGTAGCACAACG

Uppercase sequences are segments that anneal to the template DNA URAcass short/pUC19; the lowercase sequences are 50 nt upstream (ML080), resp. downstream (ML081) of the target ORF.



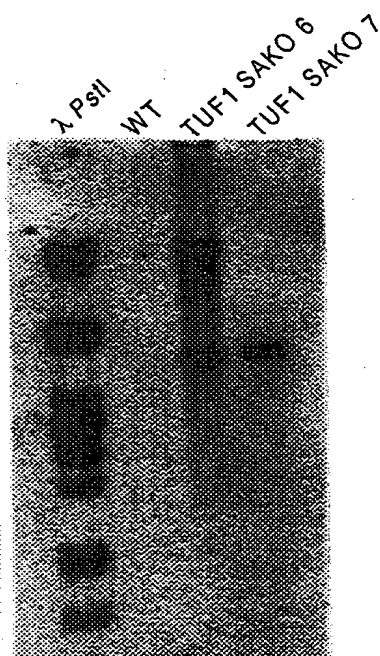
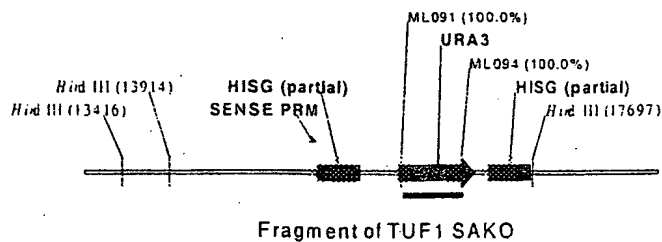
Amplification of STH fragments for TUF1 (ML080/ML094 \Rightarrow TUF1_STH-5'URA3 and ML081/ML092 \Rightarrow TUF1_STH-3'URA3). TUF1_STH-5'URA3 has a 50 bp terminal homology region upstream of the TUF1 ORF and a 3' incomplete URA3 marker; while TUF1_STH-3'URA3 has a 50 bp terminal homology region downstream of the TUF1 ORF and a 5' incomplete URA3 marker. *In vivo*, only an intact URAblast cassette can be formed when recombination occurs between the overlapping truncated URA3 sequences of the respective STH fragments.



Southern results:

FIG. 31 (CONTINUED 1).

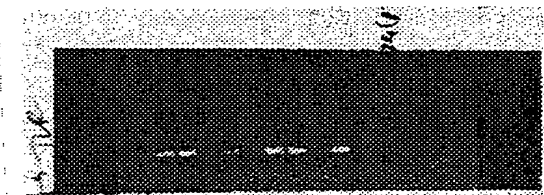
Presentation of disrupted allele



HindIII digest
URA3 probe
Expected band: 3783 bp

PCR results:

- PCR analysis was performed using the primer combination ML091/ML093 (see figure), amplifying a URA3 fragment. **Band of 755 bp points to correct homologous recombination of URA3 overlapping fragments.** TUF1 SAKO 7 is clearly positive (SAKO stands for single allele knock out).

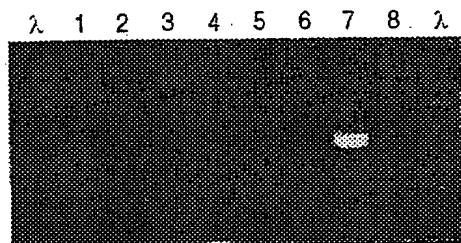


lane 1: λ PstII
lane 2 to 9: potential TUF1 SAKOs 1 to 8
lane 10: Uracass short/pUC19 (positive control)
lane 11: water
lane 12: CAI4

51/53

FIG. 31(CONTINUED 2).

- To check correct integration into the genome, PCR was performed with primer sets ML090/ML097. For TUF1 SAKO 7 a clear signal was obtained of the correct length of 1825 bp.



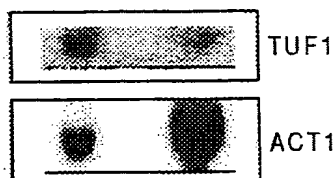
λ: λ PstI

1 to 8: potential TUF1 SAKOs 1 to 8

Northern analysis:

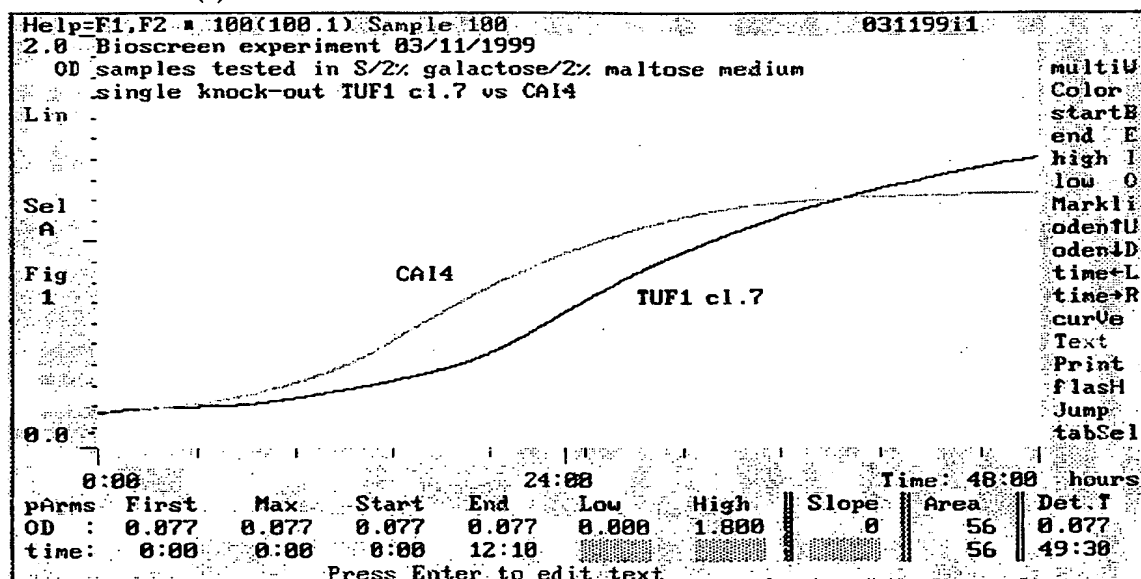
88% inhibition

TUF1 SAKO 7



Growth curve(s) :

Date : 03/11/1999



HTS screen :

FIG. 31 (CONTINUED 3)

Bioscreen test of TUF1 knock-out clone 7 and WT growth in presence of hygromycin B

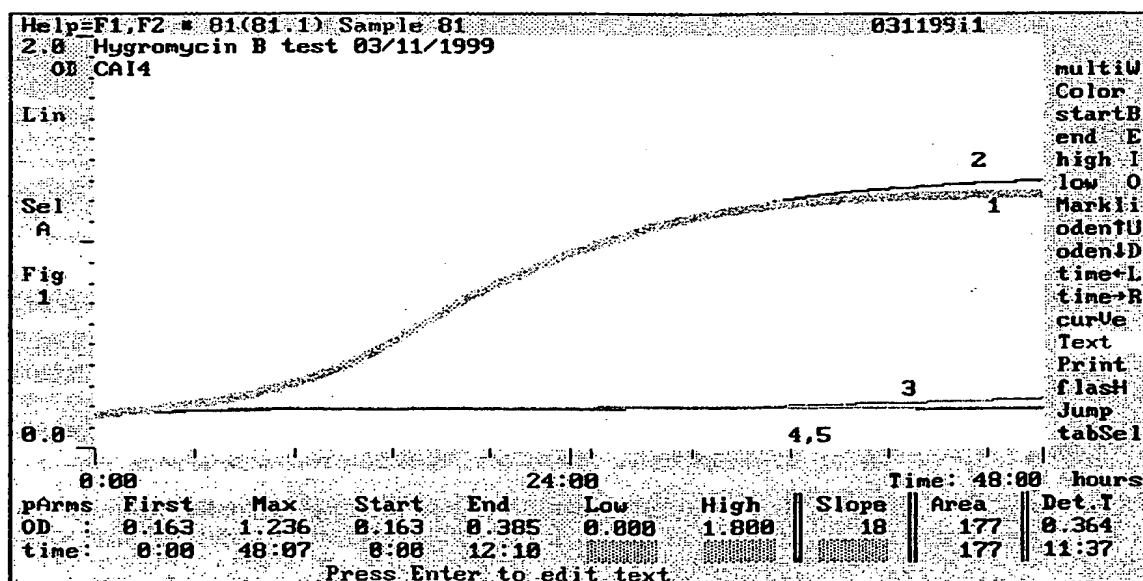
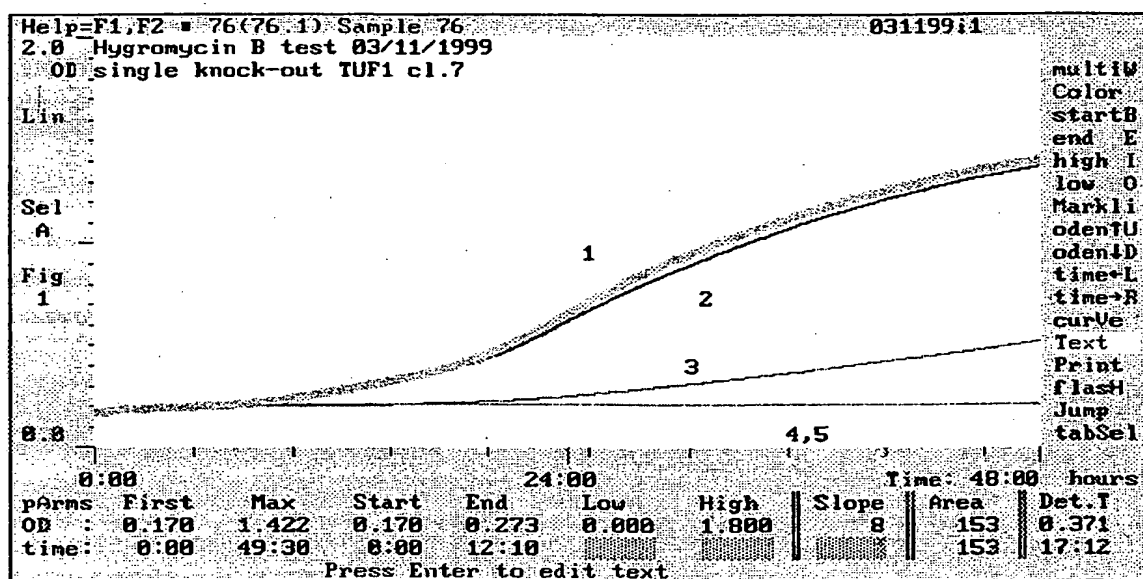
dilutions prepared

Stock solution of 53 mg/ml was prepared for hygromycin B.

From this solution dilutions were prepared of:

4000µg/ml, 3000µg/ml, 2000µg/ml and 1000µg/ml

Growth curves for TUF1 knock-out and WT in the presence of hygromycin B



Legend:

*53/53**FIG. 31 (CONTINUED 4).*

- 1: S/2% gal/2% mal medium containing 0 $\mu\text{g/ml}$ Hygromycin B
- 2: S/2% gal/2% mal medium containing 1000 $\mu\text{g/ml}$ Hygromycin B
- 3: S/2% gal/2% mal medium containing 2000 $\mu\text{g/ml}$ Hygromycin B
- 4: S/2% gal/2% mal medium containing 3000 $\mu\text{g/ml}$ Hygromycin B
- 5: S/2% gal/2% mal medium containing 4000 $\mu\text{g/ml}$ Hygromycin B

SEQUENCE LISTING

<110> Janssen Pharmaceutica N. V.

<120> Drug Targets in Candida Albicans

<130> 53731/000

<140>

<141>

<150> 982204122.0

<151> 1998-12-04

<160> 16

<170> PatentIn Ver. 2.0

<210> 1

<211> 438

<212> DNA

<213> Candida albicans

<400> 1

```
aacagctggt cttctgctaa tacattcaac cctttccata tctatactcc aacaatatga 60
taactgatga acaattgaat accattgcat tgacatttgg ttttgcttca ataataattaa 120
tcataatata tcatgccata tctactaatg tacataaatt agaagatgaa accccatcat 180
cttcattttac cagaacaaat actactgaaa ctactgttgc aagtaagaaa aagaagtaat 240
aactgatgga tttttcttcc taccaccaat tgaataatgc tagacttggt ggtgtgctac 300
aaatatattca aaagaaaata cgaatacttt ataaaatggt aagaacggaa gatgggtttct 360
catttataca ctaaatacaa atcacatata catacacaaa cacaatatata tacatacacc 420
tatatccctt tatttgat 438
```

<210> 2

<211> 1380

<212> DNA

<213> Candida albicans

<400> 2

```
atgttaaaaa cactaactca aactttacgc ttaactggga aagctttccc aaaggtccgt 60
ccggccttga tcagaacctc cgctgccttc gaccgttcta aacctcatgt caacattggt 120
actattggtc atgttgatca tggtaaaact acattgactg ctgctatcac caaagtttta 180
gccgaacaag gtggtgccaa cttcttggat tatggttcta ttgatagagc tccagaagaa 240
agagctagag gtatcactat ttccactgcc cacgttgaat acgaaaccaa gaacagacac 300
tatgccacg ttgattgtcc aggacacgct gattatatca aaaatatgat tactggtgcc 360
gctcaaatgg atggtgctat cattgttggt gctgccactg atggtcaaata gcctcaaacc 420
agagaacatt tgttattggc cagacaagtt ggtgttcaag acctgggtgt gtttgtcaac 480
aaagtcgata ctattgatga ccctgaaatg ttggaattag tcgaaatgga aatgagagaa 540
```

ttgttatcca cctacgggtt tgatggtgac aacactccag ttattatggg atctgcttta 600
 atggcttttg aagacaagaa accagaaatt ggtaagggaag ctatcttgaa attgttagat 660
 gctgtcgatg aacacattcc aactccatca agagacttgg aacaaccatt tttgttacca 720
 gttgaagacg tgttctccat ctccggtaga ggaactgttg tcaactggtg agttgaaaga 780
 ggtgttttga agaagggtga agaaatcgaa attgttggtg gttttgacaa accttacaag 840
 actactgtta ccggtattga aatgttcaaa aaagaattag actctgctat ggctggtgac 900
 aactgtggtg ttttgttaag aggtgttaaa agagatgaaa tcaagagagg tatggttttg 960
 gccaaaccag gtactgctac ttctcacaag aagttcttgg ctcccttgta tttttgact 1020
 tccgaagaag gtggtcggtc cactccattt ggtgaagggt acaagcctca atgcttcttc 1080
 agaactaacg atgtcactac cacattttca ttcccagaag gagaagggtg tgatcattct 1140
 caaatgatca tgccagggtg caacattgaa atggttggtg aattgatcaa atcttgcca 1200
 ttagaagtca accaacgttt caacttgaga gaagggtgta aaactgttgg tactggtttg 1260
 attaccagaa tcacgaata aacagaatgt gcactgtgaa taataaaaag aaaagaggta 1320
 tatatagggtg actttgtatt ttgtattgaa caataaaatt ctgtaaatag taagggcctc 1380

<210> 3

<211> 2283

<212> DNA

<213> *Candida albicans*

<400> 3

gaattcgccc ttaagcactc gtttcaacta tacattcagt aacaacaccc ttaatttacc 60
 aaactacatt aatggaagta acacaacgga cgcagagtca gacacaacca acacaacagt 120
 caccgacaac tcagacgcaa acccaaagca aagaggacca gaataggatt tgtcaattga 180
 tttgctccac gggtcagttt ggcaattatg atttgaatat caacgataaa actatcgta 240
 aaggtaaaat gacgtggtat tttggaagag accccaactc agatttgcaa gtggcgctcg 300
 cgctcgagaat ttcaaacaag ctttttcaaa tctgggtcaa cttcaatgat aaatcactat 360
 ggataaagga cacttcaact aacgggacac accttaacaa cagtcgattg gtgaaaggat 420
 caaactacct tcttaatcag ggtgatgaaa tagcagtagg ggttggtaga gacgaggacg 480
 ttgtgagggt tgctggtgtc tttggtgaca aatacaaccc ggcaaagcta cctgattcga 540
 ccaacacaat taaagatgaa ggaatatata aagactttat tgtgaaaaat gaaacgatag 600
 gccaaaggagc atttgccact gtgaaaaagg cgattgaacg atctacgggc gagtcgtacg 660
 cgggtgaagat tataaatcga agaaaagcat taaataccgg tgggtggaagt gccatggcag 720
 gagtggaccg tgaattgtcc atattagagc ggctcaacca cccaaatata gttgctctaa 780
 aagcttttta tgaagatatg gacaattact atattgtgat ggaattggtg ccgggcgggtg 840
 atttgatgga ctttgtggct gcaaacggtg caataggaga agacgcaaca caagtgatca 900
 cgaaacagat tctagaagga attgcctatg ttcataattt aggaatctcc catcgtgatt 960
 tgaagccaga taatattttg attatgcaag atgacccaat acttggttaa atcaccgact 1020
 ttggattggc aaaattcagt gacaatctga cgtttatgaa aactttttgt ggtacattgg 1080
 cgtatgttgc tcccgaagtt atcaccggta agtatggatc atcgcagatg gaactgcaac 1140
 aaaaggacaa ctactcttcc ttgggtgaca tttggtcttt gggatgtttg gtttatgtac 1200
 ttttaacttc tcatattacca ttcaacggga aaaaccagca acaaatgttt gccaaagatca 1260
 aaaggggcga atttcatgag gctccattaa attcatacga cttttctgaa gacggaagag 1320
 atttcttgca gtgctgccta cagggttaatc ctaaaactaag gatgacggct gctgaagctt 1380
 tgaaacataa atggttgcaa gacttgatg aagaggattc tgtcaaatca ttgagtttat 1440
 cgcaatcaca gtcgcaacaa tctcgaaaga tagataatgg tatccatata gaatcattga 1500
 gcaaaattga tgaagacgtt atgcttcgtc cattggatag cgaaagaaat aggaaatcaa 1560
 gtaaacagca agatttcaag gtacccaagc gtgtgattcc gttatctcaa catcctgcaa 1620

```

caccgttacc aatgtcaca cggaaaaaga ggccgtatca aatagaccct agaacaaca 1680
aaaaagtcga tttggaagaa cctctgacaa gcaagaaagt caagctaagt gattccgttg 1740
ttgcggaaga ctacttgaag ttggggccac ttgcaaattc gttattccaa gaaacaataa 1800
atatttcaaa gtccccgttt tctttcgga gaaatgacac ttgtgattgc gagatagacg 1860
acgacagact atccaaactt catttgtgtca ttaccaaaaga aaacgactct atatgggttat 1920
tggataagag tactaactcg tgcttgggtca acaatactag tgttggaaaa ggcaacaaag 1980
ttttgcttag aggaggggag atattacatc tcttctttga cccattgtca ctgcaacata 2040
taggtttcaa agtagtcctt gttgatcaac tgtctgggtga acataagagt caagtggagg 2100
ttttgaaaca aacctcagaa gaaatgaata ttattccact tatttctggt ttaagtagta 2160
taagttcata gatttagcat atatacaagc atttcctata gaaacaaagg ttcattaatt 2220
tagttattta cctccatgca attacattta cttcttcttc caagggcgaa ttctgcagat 2280
atc 2283

```

<210> 4

<211> 826

<212> DNA

<213> *Candida albicans*

<400> 4

```

atgggtagta tgtgaagata caatattgaa agtgtttact agaatatcta agatgtttga 60
gcccattggac atttttggat ttgataatta aaaaaagtag caatagatta ttgcgttgga 120
gaaagaatca ccatagttgc aagatttgat agatgttaaa atgttcacgc aggcgaaaga 180
tgtaacatct cttaaagtaa gaagaatatg gacatgaata aaaatagata gcactatttt 240
ggaacttggt gaagatatta aaatagaatg ggatttcaac atagatattc aaagtaacga 300
aacctcacia tcaaataaaa acaacagtaa tactaacaat tcaattttta tttttataga 360
gggtactcca tctttaggta aacgtcacia caaatctcac accttatgta acagatgtgg 420
ccgtcgttca ttccacgtcc aaaagaagac ctgttcttct tgtgggtacc cagctgctaa 480
aatgagatct cacaactggg ctttaaaagc caaaagaaga agaactactg gtaccggtag 540
aatggcttac ttgaaacacg ttaccagaag attcaagaac ggtttccaaa ctgggtgtgc 600
taaagctcaa accccttccg cttaaaactaa ttactgaagt tattgggtcat gcattagtca 660
ttattcatta aagtcattgt aagcatagca aaggaagaat tgggttagatt cttgtttaaa 720
atgtaatgac tatttaatat ctgtttaaat aagaggttta gtctttattt ttttacgtat 780
acaccaaaaa aaaaagaaac aaataaaatc tgtatattaa tgttgg 826

```

<210> 5

<211> 978

<212> DNA

<213> *Candida albicans*

<400> 5

```

atgggtacta gtacaagtga agcattgaag aacatcaaaa acaaacagcg aagacagaaa 60
gtttttgcag aaataaaaca tgaaaagaat aaacaacgtc ataagcaaag agccgaaaga 120
gctaagggaag aaagagaaaa ccccgaaata agagaggaaa gaatagcagc taatatccca 180
gatactatag atagcaaacg tatttatgat gagactatag ctgctgaagt tgaaggagat 240
gacgagtttc agtcataatt cactaacttg ttggaagaac caaagatttt gttgacaaca 300
agtgcgaatg ctaaaaaacc ggcctatgaa tttgcagaca tgatcatgga ctttttaccg 360
aatgtgacat ttatcaaaaag gaagaaggaa tatacaatgc aagatatggc caaatattgc 420
tcgaatagag acttcaactgc attgcttgtc atcaacgaag acaagaagaa ggtcaatggg 480

```

```

ataacgctca tcaatttacc tgaagggcca acatttttatt tttcgattac atcaatagtt 540
gatgggaaaa gaattaaggg acacgggaaa gctgggtgatt atttacctga gattgtattg 600
aataatttca attcaagatt gggtaaaaact gtgggaagac tatttcaaag tattttccct 660
cataaacctg aacttcaagg aagacaagtg attactttgc acaatcaacg tgattatatt 720
tttttcagaa gacatagata tattttcaga aatgaggaaa aggttggatt gcaggaattg 780
ggtccgcagt ttacattaaa gctaagaaga atgcaaaaagg gagtacgtgg tgatgttggt 840
tggaacaca gaccagatat ggaaagagat aagaagaagt tttatttata agcgggtgta 900
taaaggtagt agtagtgcgt ttataagtat gtgtgtgtgt ttatgcatag atgtgtaaag 960
agtaatacag ctaattcg                                     978

```

<210> 6

<211> 619

<212> DNA

<213> *Candida albicans*

<400> 6

```

aactaatttg tttaaacatc aataccaaga agttttttaca attcaatccc acatacacca 60
ttaattatga attctgaaaa gattattgaa gttatcattg ctattttctt accaccagta 120
gctgtgttta tgaaatgtgg tgccactacc ccattatgga ttaacttggg attatgtatc 180
tttatttggg tccctgctat cttacatgcc ttatacgttg tgttgaaaga ttaaacaac 240
accagagatt tactgcttga tgaattgatt actccaaaga gttgtgacta gttcccagtg 300
tgtttttttt gccttccaac tttcttttac atttttccat tactaccact gtcttcccc 360
ctattttgca gagttttcaa aatttatcca aaacatgtta gtcattaaac catattatta 420
taattattct tttttgtatt tttttccctt aaaacacgtt aatttattaa tcgtttcgtt 480
gtttggtatt ttattttttt gtatttatca attggaatat atatctatac atgaatttat 540
tatccattgt accaattggt aaaacatttt gtagtttttt tgttactagt ataaaannat 600
aataaaagtt tanttcaac                                     619

```

<210> 7

<211> 2319

<212> DNA

<213> *Candida albicans*

<400> 7

```

atgacattag ggttcgataa attcataagc aaggtcagca ctcatagacg tcaatctgaa 60
ccatcaatct tggaaatcgc agccaccaat tctcaaaata aatcgagaag gctaagtatg 120
gataatggtc attgttatgt tcgtgaatca actaataatc atcatcattt aaataccgtc 180
gttgataatt tacgacagcg tgcgggatcg ttttcattta tttcacatca ccataatcac 240
catcagaata gtcacgataa ttatactgtc gatcccccta catcaaacgg agcacgaatt 300
tcccgatcac gttcacgttc caaatcagtt gggcacggag aagcaatatc accagcgtat 360
ttttccaaga ataaaaccaa agatttagtg aaacaggaaa cagcacatat cattctgaag 420
aaattactca acatgttaca agatttggat ttacaaaacc ctattgcatt gaaaacaata 480
tcacaaggtt cagaatcaaa gttttgtaaa atctacgtgt ctaacactaa taattgtatt 540
tacttaccag cagcaagttc aacaagtttc acttatgaag atgatgaaaa tggcggcggt 600
ataattgctg aagatagaaa tgatgaaatg ccaacagcag ttaataacaa tactttgtca 660
atggatagta taaatcattc agagactgat ttcttggtt ctccaccacc tccagattta 720
ttttctaaaa tgaaatcatt ccattcacca aattacttga cttcaaaaat cgattctgaa 780
tgtccaattc cacatacatt tgctgtgatt gttgaattaa ccaaggactc tttgattatt 840

```

```

aaagatcttc atttccaatt tcagtcatta actaccattt tatggccaac tgggggatgca 900
tataatcgga ctcatgccaa ggagaaattt accattggga atatggaatg gcgtacatct 960
ttaagcgacg ccgactatta tatcaatagt tctaattcca acgatgttaa gctgaaaaac 1020
ttgggtcctg aagatcttat taatcgaact agagaataca aattaatcga tattgaagaa 1080
ccaaacaatt catcaaacag tttactggat gatgacatgg atattaataa tattacgtcg 1140
ccattatcaa cgtcaccaac atcaagttca acttcaacaa attcaacctc caactcattg 1200
ggttcagatt catataaagc tgggtctttat gtattttttat taccaatctt attgccagaa 1260
catattcctg cttccattgt ttctattaat gggttcattgg ctcatacatt actggttgaa 1320
tgcaataaat atactgataa gttgaatcgg aaatcaaaaag tatcagcatc gtacaattta 1380
cctatgggtcc gtactccacc aaacattgggt aattccattg ctgataagcc aatttatgtt 1440
aataggattt ggaatgatgc cgtacattat attataactt tcccccgcaa atatgttact 1500
ttgggttggtg aacacatgat aaatgtgaaa ttactgcca tggtgaaaga tgtggttatc 1560
aagcgtatta aatttaattg attggagaga ataacttatg tttccaaaaa tttatcacga 1620
gaatatgatt atgatagtga agacccttat tgtattcatc cagtttctaa agaaaataaa 1680
gtacgtgaac gtgttggtgc gttatatgaa ttgaaaacga aggcaaaaaca atcttctggg 1740
ggacatcttg aagcttataa acaagaagtt atgaaatgtc cggaaaataa ctttttattt 1800
tcttgttatg aggttgaaaa tgataataat aacggcaacg gcaacggcaa cggcaacgga 1860
aacaagaacg ttaaacaaaa gaataaagat caaccaatga ttgctacacc tttagatata 1920
aatgtttctt taccattttt aactactatg tctgatagtt taattatgac atcagccata 1980
gaagaagaag gttcagatct gctcatata tcaagaagag ggtcggcagt gagtatgact 2040
gataataata ctaccccaag taacaataac cctttatctc catttttggg agcagtggaa 2100
actaatgggtg ctagtataaa tgaaattggg gatcatacat tattccctga ttctaatttt 2160
cgacatattg aaattaaaca tcgattacaa gttacattta ggatttctaa accggatctg 2220
gataataaaa tgcataatta tgaagtgggt attgataccc ccatcgtttt acttagttca 2280
aatgtcaag aagattctcc tcttccttat agttctgta 2319

```

<210> 8

<211> 255

<212> DNA

<213> Candida albicans

<400> 8

```

aacgttcgtg caaaaggcta tactggtgat atccacgcag atgaagagca agtttaataca 60
actctttgtc aattaatgct gtacttggtt tcatTTTTatt tgctggcatt taaagaatac 120
ccatagttca gaaaataaaa ttgaaaaatt taaaaaaaaa cgcaatatca ttcatttttt 180
ttgttttttt gacaataata ttaatatgta gttaccaatg ttttttagatt ttatatgttt 240
tgaaaaaata gtttg

```

<210> 9

<211> 119

<212> DNA

<213> Candida albicans

<400> 9

```

aaccttaciaa tcattatacc aactatcaaa atcataagac tcttnaactt ctgttttttga 60
tagttgggtat aatgattta: gattatctt aattcattat tattagtttc ggtcacaaa 119

```

<210> 10

<211> 60
 <212> PRT
 <213> Candida albicans

<400> 10
 Met Ile Thr Asp Glu Gln Leu Asn Thr Ile Ala Leu Thr Phe Gly Phe
 1 5 10 15
 Ala Ser Ile Ile Leu Ile Ile Ile Tyr His Ala Ile Ser Thr Asn Val
 20 25 30
 His Lys Leu Glu Asp Glu Thr Pro Ser Ser Ser Phe Thr Arg Thr Asn
 35 40 45
 Thr Thr Glu Thr Thr Val Ala Ser Lys Lys Lys Lys
 50 55 60

<210> 11
 <211> 426
 <212> PRT
 <213> Candida albicans

<400> 11
 Met Leu Lys Thr Leu Thr Gln Thr Leu Arg Leu Thr Gly Lys Ala Phe
 1 5 10 15
 Pro Lys Val Arg Pro Ala Leu Ile Arg Thr Tyr Ala Ala Phe Asp Arg
 20 25 30
 Ser Lys Pro His Val Asn Ile Gly Thr Ile Gly His Val Asp His Gly
 35 40 45
 Lys Thr Thr Leu Thr Ala Ala Ile Thr Lys Val Leu Ala Glu Gln Gly
 50 55 60
 Gly Ala Asn Phe Leu Asp Tyr Gly Ser Ile Asp Arg Ala Pro Glu Glu
 65 70 75 80
 Arg Ala Arg Gly Ile Thr Ile Ser Thr Ala His Val Glu Tyr Glu Thr
 85 90 95
 Lys Asn Arg His Tyr Ala His Val Asp Cys Pro Gly His Ala Asp Tyr
 100 105 110
 Ile Lys Asn Met Ile Thr Gly Ala Ala Gln Met Asp Gly Ala Ile Ile
 115 120 125

Val Val Ala Ala Thr Asp Gly Gln Met Pro Gln Thr Arg Glu His Leu
 130 135 140

Leu Leu Ala Arg Gln Val Gly Val Gln Asp Leu Val Val Phe Val Asn
 145 150 155 160

Lys Val Asp Thr Ile Asp Asp Pro Glu Met Leu Glu Leu Val Glu Met
 165 170 175

Glu Met Arg Glu Leu Leu Ser Thr Tyr Gly Phe Asp Gly Asp Asn Thr
 180 185 190

Pro Val Ile Met Gly Ser Ala Leu Met Ala Leu Glu Asp Lys Lys Pro
 195 200 205

Glu Ile Gly Lys Glu Ala Ile Leu Lys Leu Leu Asp Ala Val Asp Glu
 210 215 220

His Ile Pro Thr Pro Ser Arg Asp Leu Glu Gln Pro Phe Leu Leu Pro
 225 230 235 240

Val Glu Asp Val Phe Ser Ile Ser Gly Arg Gly Thr Val Val Thr Gly
 245 250 255

Arg Val Glu Arg Gly Val Leu Lys Lys Gly Glu Glu Ile Glu Ile Val
 260 265 270

Gly Gly Phe Asp Lys Pro Tyr Lys Thr Thr Val Thr Gly Ile Glu Met
 275 280 285

Phe Lys Lys Glu Leu Asp Ser Ala Met Ala Gly Asp Asn Cys Gly Val
 290 295 300

Leu Leu Arg Gly Val Lys Arg Asp Glu Ile Lys Arg Gly Met Val Leu
 305 310 315 320

Ala Lys Pro Gly Thr Ala Thr Ser His Lys Lys Phe Leu Ala Ser Leu
 325 330 335

Tyr Ile Leu Thr Ser Glu Glu Gly Gly Arg Ser Thr Pro Phe Gly Glu
 340 345 350

Gly Tyr Lys Pro Gln Cys Phe Phe Arg Thr Asn Asp Val Thr Thr Thr
 355 360 365

Phe Ser Phe Pro Glu Gly Glu Gly Val Asp His Ser Gln Met Ile Met
 370 375 380

Pro Gly Asp Asn Ile Glu Met Val Gly Glu Leu Ile Lys Ser Cys Pro
 385 390 395 400

Leu Glu Val Asn Gln Arg Phe Asn Leu Arg Glu Gly Gly Lys Thr Val
 405 410 415

Gly Thr Gly Leu Ile Thr Arg Ile Ile Glu
 420 425

<210> 12

<211> 699

<212> PRT

<213> Candida albicans

<400> 12

Met Glu Val Thr Gln Arg Thr Gln Ser Gln Thr Gln Pro Thr Gln Gln
 1 5 10 15

Ser Pro Thr Thr Gln Thr Gln Thr Gln Ser Lys Glu Asp Gln Asn Arg
 20 25 30

Ile Cys Gln Leu Ile Cys Ser Thr Gly Gln Phe Gly Asn Tyr Asp Leu
 35 40 45

Asn Ile Asn Asp Lys Thr Ile Val Gln Gly Lys Met Thr Trp Tyr Phe
 50 55 60

Gly Arg Asp Pro Asn Ser Asp Leu Gln Val Ala Ser Ser Ser Arg Ile
 65 70 75 80

Ser Asn Lys His Phe Gln Ile Trp Leu Asn Phe Asn Asp Lys Ser Leu
 85 90 95

Trp Ile Lys Asp Thr Ser Thr Asn Gly Thr His Leu Asn Asn Ser Arg
 100 105 110

Leu Val Lys Gly Ser Asn Tyr Leu Leu Asn Gln Gly Asp Glu Ile Ala
 115 120 125

Val Gly Val Gly Arg Asp Glu Asp Val Val Arg Phe Val Val Val Phe
 130 135 140

Gly Asp Lys Tyr Asn Pro Ala Lys Leu Pro Asp Ser Thr Asn Thr Ile
 145 150 155 160

Lys Asp Glu Gly Ile Tyr Lys Asp Phe Ile Val Lys Asn Glu Thr Ile
 165 170 175

Gly Gln Gly Ala Phe Ala Thr Val Lys Lys Ala Ile Glu Arg Ser Thr
 180 185 190

Gly Glu Ser Tyr Ala Val Lys Ile Ile Asn Arg Arg Lys Ala Leu Asn
 195 200 205

Thr Gly Gly Gly Ser Ala Met Ala Gly Val Asp Arg Glu Leu Ser Ile
 210 215 220

Leu Glu Arg Leu Asn His Pro Asn Ile Val Ala Leu Lys Ala Phe Tyr
 225 230 235 240

Glu Asp Met Asp Asn Tyr Tyr Ile Val Met Glu Leu Val Pro Gly Gly
 245 250 255

Asp Leu Met Asp Phe Val Ala Ala Asn Gly Ala Ile Gly Glu Asp Ala
 260 265 270

Thr Gln Val Ile Thr Lys Gln Ile Leu Glu Gly Ile Ala Tyr Val His
 275 280 285

Asn Leu Gly Ile Ser His Arg Asp Leu Lys Pro Asp Asn Ile Leu Ile
 290 295 300

Met Gln Asp Asp Pro Ile Leu Val Lys Ile Thr Asp Phe Gly Leu Ala
 305 310 315 320

Lys Phe Ser Asp Asn Ser Thr Phe Met Lys Thr Phe Cys Gly Thr Leu
 325 330 335

Ala Tyr Val Ala Pro Glu Val Ile Thr Gly Lys Tyr Gly Ser Ser Gln
 340 345 350

Met Glu Ser Gln Gln Lys Asp Asn Tyr Ser Ser Leu Val Asp Ile Trp
 355 360 365

Ser Leu Gly Cys Leu Val Tyr Val Leu Leu Thr Ser His Leu Pro Phe
 370 375 380

Asn Gly Lys Asn Gln Gln Gln Met Phe Ala Lys Ile Lys Arg Gly Glu
 385 390 395 400

Phe His Glu Ala Pro Leu Asn Ser Tyr Asp Ile Ser Glu Asp Gly Arg
 405 410 415

Asp Phe Leu Gln Cys Cys Leu Gln Val Asn Pro Lys Leu Arg Met Thr
 420 425 430

Ala Ala Glu Ala Leu Lys His Lys Trp Leu Gln Asp Leu Tyr Glu Glu
 435 440 445

Asp Ser Val Lys Ser Leu Ser Leu Ser Gln Ser Gln Ser Gln Gln Ser
 450 455 460

Arg Lys Ile Asp Asn Gly Ile His Ile Glu Ser Leu Ser Lys Ile Asp
 465 470 475 480

Glu Asp Val Met Leu Arg Pro Leu Asp Ser Glu Arg Asn Arg Lys Ser
 485 490 495

Ser Lys Gln Gln Asp Phe Lys Val Pro Lys Arg Val Ile Pro Leu Ser
 500 505 510

Gln His Pro Ala Thr Pro Leu Pro Met Ser Gln Pro Lys Lys Arg Pro
 515 520 525

Tyr Gln Ile Asp Pro Arg Thr Asn Lys Lys Val Asp Leu Glu Glu Pro
 530 535 540

Ser Thr Ser Lys Lys Val Lys Leu Ser Asp Ser Val Val Ala Glu Asp
 545 550 555 560

Tyr Leu Lys Leu Gly Pro Leu Ala Asn Ser Leu Phe Gln Glu Thr Ile
 565 570 575

Asn Ile Ser Lys Ser Pro Phe Ser Phe Gly Arg Asn Asp Thr Cys Asp
 580 585 590

Cys Glu Ile Asp Asp Asp Arg Leu Ser Lys Leu His Cys Val Ile Thr
 595 600 605

Lys Glu Asn Asp Ser Ile Trp Leu Leu Asp Lys Ser Thr Asn Ser Cys
 610 615 620

Leu Val Asn Asn Thr Ser Val Gly Lys Gly Asn Lys Val Leu Leu Arg
 625 630 635 640

Gly Gly Glu Ile Leu His Leu Phe Phe Asp Pro Leu Ser Ser Gln His
 645 650 655

Ile Gly Phe Lys Val Val Leu Val Asp Gln Ser Ser Gly Glu His Lys
 660 665 670

Ser Gln Val Glu Val Leu Lys Gln Thr Ser Glu Glu Met Asn Ile Ile
 675 680 685

Pro Leu Ile Ser Gly Leu Ser Ser Ile Ser Ser
690 695

<210> 13

<211> 295

<212> PRT

<213> Candida albicans

<400> 13

Met Gly Thr Ser Thr Ser Glu Ala Leu Lys Asn Ile Lys Asn Lys Gln
1 5 10 15

Arg Arg Gln Lys Val Phe Ala Glu Ile Lys His Glu Lys Asn Lys Gln
20 25 30

Arg His Lys Gln Arg Ala Glu Arg Ala Lys Glu Glu Arg Glu Asn Pro
35 40 45

Glu Leu Arg Glu Glu Arg Ile Ala Ala Asn Ile Pro Asp Thr Ile Asp
50 55 60

Ser Lys Arg Ile Tyr Asp Glu Thr Ile Ala Ala Glu Val Glu Gly Asp
65 70 75 80

Asp Glu Phe Gln Ser Tyr Phe Thr Asn Leu Leu Glu Glu Pro Lys Ile
85 90 95

Leu Leu Thr Thr Ser Ala Asn Ala Lys Lys Pro Ala Tyr Glu Phe Ala
100 105 110

Asp Met Ile Met Asp Phe Leu Pro Asn Val Thr Phe Ile Lys Arg Lys
115 120 125

Lys Glu Tyr Thr Met Gln Asp Met Ala Lys Tyr Cys Ser Asn Arg Asp
130 135 140

Phe Thr Ala Leu Leu Val Ile Asn Glu Asp Lys Lys Lys Val Asn Gly
145 150 155 160

Ile Thr Leu Ile Asn Leu Pro Glu Gly Pro Thr Phe Tyr Phe Ser Ile
165 170 175

Thr Ser Ile Val Asp Gly Lys Arg Ile Lys Gly His Gly Lys Ala Gly
180 185 190

Asp Tyr Leu Pro Glu Ile Val Leu Asn Asn Phe Asn Ser Arg Leu Gly

195 200 205
 Lys Thr Val Gly Arg Leu Phe Gln Ser Ile Phe Pro His Lys Pro Glu
 210 215 220
 Leu Gln Gly Arg Gln Val Ile Thr Leu His Asn Gln Arg Asp Tyr Ile
 225 230 235 240
 Phe Phe Arg Arg His Arg Tyr Ile Phe Arg Asn Glu Glu Lys Val Gly
 245 250 255
 Leu Gln Glu Gly Pro Gln Phe Thr Leu Lys Leu Arg Arg Met Gln Lys
 260 265 270
 Gly Val Arg Gly Asp Val Val Trp Glu His Arg Pro Asp Met Glu Arg
 275 280 285
 Asp Lys Lys Lys Phe Tyr Leu
 290 295

<210> 14
 <211> 55
 <212> PRT
 <213> Candida albicans

<400> 14
 Met Asn Ser Glu Lys Ile Ile Glu Val Ile Ile Ala Ile Phe Leu Pro
 1 5 10 15
 Pro Val Ala Val Phe Met Lys Cys Gly Ala Thr Thr Pro Leu Trp Ile
 20 25 30
 Asn Leu Val Leu Cys Ile Phe Ile Trp Phe Pro Ala Ile Leu His Ala
 35 40 45
 Leu Tyr Val Val Leu Lys Asp
 50 55

<210> 15
 <211> 773
 <212> PRT
 <213> Candida albicans

<400> 15
 Met Thr Leu Gly Phe Asp Lys Phe Ile Ser Lys Val Ser Thr His Arg
 1 5 10 15

Arg Gln Ser Glu Pro Ser Ile Leu Glu Ile Ala Ala Thr Asn Ser Gln
 20 25 30

Asn Lys Ser Arg Arg Leu Ser Met Asp Asn Gly His Cys Tyr Val Arg
 35 40 45

Glu Ser Thr Asn Asn His His His Leu Asn Thr Val Val Asp Asn Leu
 50 55 60

Arg Gln Arg Ala Gly Ser Phe Ser Phe Ile Ser His His His Asn His
 65 70 75 80

His Gln Asn Ser His Asp Asn Tyr Thr Val Asp Pro Leu Thr Ser Asn
 85 90 95

Gly Ala Arg Ile Ser Arg Ser Arg Ser Arg Ser Lys Ser Val Gly His
 100 105 110

Gly Glu Ala Ile Ser Pro Ala Tyr Phe Ser Lys Asn Lys Thr Lys Asp
 115 120 125

Leu Val Lys Gln Glu Thr Ala His Ile Ile Ser Lys Lys Leu Leu Asn
 130 135 140

Met Leu Gln Asp Leu Asp Leu Gln Asn Pro Ile Ala Leu Lys Thr Ile
 145 150 155 160

Ser Gln Gly Ser Glu Ser Lys Phe Cys Lys Ile Tyr Val Ser Asn Thr
 165 170 175

Asn Asn Cys Ile Tyr Leu Pro Ala Ala Ser Ser Thr Ser Phe Thr Tyr
 180 185 190

Glu Asp Asp Glu Asn Gly Gly Val Ile Ile Ala Glu Asp Arg Asn Asp
 195 200 205

Glu Met Pro Thr Ala Val Asn Asn Asn Thr Leu Ser Met Asp Ser Ile
 210 215 220

Asn His Ser Glu Thr Asp Phe Ser Asp Ser Pro Pro Pro Pro Asp Leu
 225 230 235 240

Phe Ser Lys Met Lys Ser Phe His Ser Pro Asn Tyr Leu Thr Ser Lys
 245 250 255

Ile Asp Ser Glu Cys Pro Ile Pro His Thr Phe Ala Val Ile Val Glu
 260 265 270

Leu Thr Lys Asp Ser Leu Ile Ile Lys Asp Leu His Phe Gln Phe Gln
 275 280 285
 Ser Leu Thr Thr Ile Leu Trp Pro Thr Gly Asp Ala Tyr Asn Arg Thr
 290 295 300
 His Ala Lys Glu Lys Phe Thr Ile Gly Asn Met Glu Trp Arg Thr Ser
 305 310 315 320
 Leu Ser Asp Ala Asp Tyr Tyr Ile Asn Ser Ser Asn Ser Asn Asp Val
 325 330 335
 Lys Ser Lys Asn Leu Gly Pro Glu Asp Leu Ile Asn Arg Thr Arg Glu
 340 345 350
 Tyr Lys Leu Ile Asp Ile Glu Glu Pro Asn Asn Ser Ser Asn Ser Leu
 355 360 365
 Ser Asp Asp Asp Met Asp Ile Asn Asn Ile Thr Ser Pro Leu Ser Thr
 370 375 380
 Ser Pro Thr Ser Ser Ser Thr Ser Thr Asn Ser Thr Ser Asn Ser Leu
 385 390 395 400
 Gly Ser Asp Ser Tyr Lys Ala Gly Leu Tyr Val Phe Leu Leu Pro Ile
 405 410 415
 Leu Leu Pro Glu His Ile Pro Ala Ser Ile Val Ser Ile Asn Gly Ser
 420 425 430
 Leu Ala His Thr Leu Ser Val Glu Cys Asn Lys Tyr Thr Asp Lys Leu
 435 440 445
 Asn Arg Lys Ser Lys Val Ser Ala Ser Tyr Asn Leu Pro Met Val Arg
 450 455 460
 Thr Pro Pro Asn Ile Gly Asn Ser Ile Ala Asp Lys Pro Ile Tyr Val
 465 470 475 480
 Asn Arg Ile Trp Asn Asp Ala Val His Tyr Ile Ile Thr Phe Pro Arg
 485 490 495
 Lys Tyr Val Thr Leu Gly Cys Glu His Met Ile Asn Val Lys Leu Ser
 500 505 510
 Pro Met Val Lys Asp Val Val Ile Lys Arg Ile Lys Phe Asn Val Leu
 515 520 525

Glu Arg Ile Thr Tyr Val Ser Lys Asn Leu Ser Arg Glu Tyr Asp Tyr
 530 535 540

Asp Ser Glu Asp Pro Tyr Cys Ile His Pro Val Ser Lys Glu Asn Lys
 545 550 555 560

Val Arg Glu Arg Val Val Ser Leu Tyr Glu Leu Lys Thr Lys Ala Lys
 565 570 575

Gln Ser Ser Gly Gly His Leu Glu Ala Tyr Lys Gln Glu Val Met Lys
 580 585 590

Cys Pro Glu Asn Asn Leu Leu Phe Ser Cys Tyr Glu Val Glu Asn Asp
 595 600 605

Asn Asn Asn Gly Asn Gly Asn Gly Asn Gly Asn Gly Asn Lys Asn Val
 610 615 620

Lys Gln Lys Asn Lys Asp Gln Pro Met Ile Ala Thr Pro Leu Asp Ile
 625 630 635 640

Asn Val Ser Leu Pro Phe Leu Thr Thr Met Ser Asp Ser Leu Ile Met
 645 650 655

Thr Ser Ala Ile Glu Glu Glu Gly Ser Asp Ser Pro His Thr Ser Arg
 660 665 670

Arg Gly Ser Ala Val Ser Met Thr Asp Asn Asn Thr Thr Pro Ser Asn
 675 680 685

Asn Asn Pro Leu Ser Pro Phe Leu Gly Ala Val Glu Thr Asn Gly Ala
 690 695 700

Ser Ile Asn Glu Ile Gly Asp His Thr Leu Phe Pro Asp Ser Asn Phe
 705 710 715 720

Arg His Ile Glu Ile Lys His Arg Leu Gln Val Thr Phe Arg Ile Ser
 725 730 735

Lys Pro Asp Ser Asp Asn Lys Met His His Tyr Glu Val Val Ile Asp
 740 745 750

Thr Pro Ile Val Leu Leu Ser Ser Lys Cys Gln Glu Asp Ser Pro Pro
 755 760 765

Pro Tyr Ser Ser Val
 770

<210> 16

<211> 90

<212> PRT

<213> Candida albicans

<400> 16

Met Gly Glu Gly Thr Pro Ser Leu Gly Lys Arg His Asn Lys Ser His
1 5 10 15

Thr Leu Cys Asn Arg Cys Gly Arg Arg Ser Phe His Val Gln Lys Lys
20 25 30

Thr Cys Ser Ser Cys Gly Tyr Pro Ala Ala Lys Met Arg Ser His Asn
35 40 45

Trp Ala Leu Lys Ala Lys Arg Arg Arg Thr Thr Gly Thr Gly Arg Met
50 55 60

Ala Tyr Leu Lys His Val Thr Arg Arg Phe Lys Asn Gly Phe Gln Thr
65 70 75 80

Gly Val Ala Lys Ala Gln Thr Pro Ser Ala
85 90

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
15 June 2000 (15.06.2000)

PCT

(10) International Publication Number
WO 00/34481 A3

(51) International Patent Classification?: **C12N 15/31**,
C07K 14/40, G01N 33/50, C12Q 1/68, A61K 31/7088,
38/16, C07K 16/14

University of Gent, K.L. Ledeganckstraat 35, B-9000 Gent
(BE). **VIALARD, Jorge, Eduardo** [CA/BE]; Janssen
Pharmaceutica N.V., Turnhoutseweg 30, B-2340 Beerse
(BE).

(21) International Application Number: PCT/EP99/09833

(22) International Filing Date: 6 December 1999 (06.12.1999)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
98204122.0 4 December 1998 (04.12.1998) EP

(74) Agent: **BOULT WADE TENNANT**; 27 Furnival Street,
London EC4A 1PQ (GB).

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE,
ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP,
KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD,
MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD,
SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ,
VN, YU, ZA, ZW.

(71) Applicant (*for all designated States except US*):
JANSSEN PHARMACEUTICA N.V. [BE/BE]; Turn-
houtseweg 30, B-2340 Beerse (BE).

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,
MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **CONTRERAS,**
Roland, Henri [BE/BE]; University of Gent, K.L. Lede-
ganckstraat 35, B-9000 Gent (BE). **NELISSEN, Bart**
[BE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30,
B-2340 Beerse (BE). **DE BACKER, Marianne, Denise**
[BE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30,
B-2340 Beerse (BE). **LUYTEN, Walter, Herman, Maria,**
Louis [BE/BE]; Janssen Pharmaceutica N.V., Turnhout-
seweg 30, B-2340 Beerse (BE). **VIAENE, Jasmine, Elza**
[BE/BE]; University of Gent, K.L. Ledeganck Straat 35,
B-9000 Gent (BE). **LOGGHE, Marc, George** [BE/BE];

Published:

— With international search report.

(88) Date of publication of the international search report:
30 November 2000

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: DRUG TARGETS IN *CANDIDA ALBICANS*

(57) Abstract: The present invention is concerned with the identification of genes or functional fragments thereof from *Candida albicans* which are critical for growth and cell division and which genes may be used as selective drug targets to treat *Candida albicans* associated infections. Novel nucleic acid sequences from *Candida albicans* are also provided and which encode the polypeptides which are critical for growth of *Candida albicans*. Methods for the identification of anti-fungal compounds which inhibit fungal or yeast growth are also contemplated.

WO 00/34481 A3

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 99/09833

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/31 C07K14/40 G01N33/50 C12Q1/68 A61K31/7088
A61K38/16 C07K16/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL - EMGSS12 [Online] Entry/Acc.no. B89355, 19 March 1998 (1998-03-19) ADAMS, M. D. ET AL.: "RPCI11-27L12.TPB RPCI-11 Homo sapiens genomic clone RPCI-11-27L12, genomic survey sequence." XP002139078 the whole document	33,34
A	WO 97 36925 A (SCRIPTGEN PHARM INC ;HARVARD COLLEGE (US)) 9 October 1997 (1997-10-09) the whole document	
A	WO 97 37230 A (SCRIPTGEN PHARM INC ;HARVARD COLLEGE (US)) 9 October 1997 (1997-10-09) the whole document	
	--- -/-	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *Z* document member of the same patent family

Date of the actual completion of the international search

30 May 2000

Date of mailing of the international search report

23.08.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Smalt, R

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 99/09833

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 36707 A (UNIV ROMA ;IST SUPERIORE SANITA (IT); CASSONE ANTONIO (IT); VALLE) 21 November 1996 (1996-11-21) the whole document ---	
A	DALY S ET AL: "Isolation and characterization of a gene encoding alpha-tubulin from Candida albicans" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES,GB,ELSEVIER SCIENCE PUBLISHERS, BARKING, vol. 187, no. 2, 18 March 1997 (1997-03-18), pages 151-158, XP004093273 ISSN: 0378-1119 the whole document -----	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 99/09833

Box I Observation where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 25-28
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-28 and 32-34, all partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 25-28

Present claims 25-28 relate to a compound defined by reference to a desirable characteristic or property, namely its identifyability by the method of claim 24.

The application does not provide support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search for these claims is impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: Invention 1: claims 1-28 and 32-34, all partially

Nucleic acid molecule comprising seq.ID.1 or capable of hybridizing thereto and antisense thereof, cell containing said nucleic acid, polypeptide of seq.ID.10 encoded by said nucleic acid, expression vector comprising said nucleic acid, antibody against said peptide, use of said nucleic acid or protein for preparation of medicament or pharmaceutical composition, *C. albicans* cell comprising an induced mutation in said DNA sequence, oligonucleotides comprising 10-120 nt of said nucleic acid sequence, and method for identifying compounds which modulate expression of said nucleic acid.

2. Claims: Inventions 2-9: claims 1-28 and 32-34,
all partially and as applicable

As invention 1, but limited to the respective nucleic acid sequences 2-9, and polypeptide sequences corresponding thereto in as far as they are provided (see table 1 of the description).

3. Claims: Invention 10: claim 29-31

Method for identifying DNA sequences from a cell or organism, which encode polypeptides which are critical for growth and survival for said cell or organism, comprising screening a library of nucleic acids using a vector that either integrates into the genome of said cell or organism, or that permits expression of antisense RNA, and selecting growth-impaired cells or organisms.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 99/09833

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9736925 A	09-10-1997	CA 2250129 A EP 0904289 A	09-10-1997 31-03-1999
WO 9737230 A	09-10-1997	US 5863762 A CA 2250121 A EP 0894269 A	26-01-1999 09-10-1997 03-02-1999
WO 9636707 A	21-11-1996	IT RM950314 A AU 5777696 A EP 0826040 A	18-11-1996 29-11-1996 04-03-1998